

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



LSHTM Research Online

Osman, Nadia; (2008) Measuring Persuasive Communication in two nutrition education delivery methods in rural Mozambique. PhD thesis, London School of Hygiene & Tropical Medicine. DOI: <https://doi.org/10.17037/PUBS.01520138>

Downloaded from: <https://researchonline.lshtm.ac.uk/id/eprint/1520138/>

DOI: <https://doi.org/10.17037/PUBS.01520138>

Usage Guidelines:

Please refer to usage guidelines at <https://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license. To note, 3rd party material is not necessarily covered under this license: <http://creativecommons.org/licenses/by-nc-nd/3.0/>

<https://researchonline.lshtm.ac.uk>

**MEASURING PERSUASIVE COMMUNICATION IN TWO
NUTRITION EDUCATION DELIVERY METHODS IN RURAL
MOZAMBIQUE**

NADIA OSMAN

**Thesis submitted for the degree of Doctor of Philosophy
Faculty of Medicine, University of London**

**Nutrition and Public Health Intervention Research Unit
Department of Epidemiology and Population Health
London School of Hygiene and Tropical Medicine**

November 2007

BEST COPY

AVAILABLE

Variable print quality

ABSTRACT

Background: Nutrition education programmes are frequently implemented in developing countries with the aim of improving dietary practices. Few studies have, to date, examined the various steps in such an education process in a developing country setting. This study, uses McGuire's "Communication/Persuasion" model to assess the impact among poor rural women in Mozambique of two nutrition education delivery methods: group classes only versus group classes plus individual classes.

Methods: Study participants were randomly selected from the total population of RCT taking place in rural Mozambique: the Towards Sustainable Nutrition Improvement project. Participants were female, had a child under 5 years of age and had been randomised to receive the nutrition education program. McGuire's output steps (exposure, attention, understanding, memorisation/recall, skills acquisition, and attitude change) were assessed for both nutrition education delivery methods using a range of tools developed by the investigator, ranging from direct observations to questionnaires. The sample size varied for each output step depending on pre-defined criteria.

Results: The methods devised in the current study for the assessment of McGuire's output steps were able to detect a good degree of variability in response. Mothers receiving both the group and individual classes (intervention II) scored higher in most output steps than those receiving only the group classes (intervention I), even after adjusting for possible confounding influences. For example, those mothers in intervention II had significantly higher skills acquisition scores than those mothers in intervention I (difference of means 1.03, independent t-test, $p < 0.001$). Within the sample of intervention II mothers, the scores from measurements taken in individual classes were significantly higher than those taken in the group classes. For example, while there was no difference in observed attention scores between intervention II and intervention I mothers when assessed at group classes, observed attention scores at individual classes were significantly higher (for intervention II mothers) than those of intervention I mothers at group classes (Independent t-test, $p = 0.031$). Further analysis appeared to suggest that this advantage was due to the added benefit of receiving individual classes in addition to group classes, although this cannot be confirmed in the current study given the lack of a study arm providing only individual classes.

Conclusions: This study has for the first time used McGuire's model to systematically investigate and compare the communication process in two nutrition education delivery methods in a developing country and attempted to develop tools to measure McGuire's output steps in such a setting. McGuire's model emerged as a good conceptual framework to use for this purpose as it is practical, and allowed for the breakdown of the process into discrete steps that permitted comparisons to be made between two delivery methods. The analysis suggested that there was an advantage in providing individual classes in addition to group classes, although this clearly has important resource implications.

ACKNOWLEDGEMENTS

First I would like to thank my supervisor, Alan Dangour, for his support, guidance, advice and especially, for making time for me through out my PhD, and as well thank Saul Morris for guiding me in the first 6 months of my PhD.

I am very grateful to my colleagues, my office mates, and my friends – Laura Howe, Bilal Iqbal, Sophie Plagerson, Abigail Perry, Irma Ismail, Karen Devries, Amira Shaheen, and Liandre VanDerMerwe. Thank you for making room/office 108 so much like home. Without their support, motivation and friendship, these last 3-4 years would have certainly not been so much fun and joy. A special thanks to Laura Howe, for her patience and willingness to answer all my infinite questions and for proofreading and commenting on my thesis' chapters.

I am also grateful to Jan Low for her support during my field work, making it easier for me, and for her willingness to contribute to my work. Another special thanks for the TSNi staff, especially the health promoters, for their help and support during my field work in Mozambique.

A big thanks for the statistical support from Saowalak Hunnangkul, Tom Marshall and again from Bilal Iqbal and Laura Howe.

A big thanks for the advice from Professor George Gaskell and Professor John Elder.

Last and not least, a big thanks to my dad, not only for the financial support, but mostly for believing in me. Another big thanks for my mother for her worrying and caring. Many other thanks for the support and love received from my sister, the rest of my family and my great friends in London, Portugal, and of course, Mozambique.

TABLE OF CONTENTS

LIST OF TABLES	11
LIST OF FIGURES	13
LIST OF APPENDICES	14
CHAPTER 1 --- INTRODUCTION	15
MALNUTRITION AND NUTRITION EDUCATION: BACKGROUND OF STUDY	15
1.1 Overview and purpose of the study	15
1.2 Contents	17
1.3 Background of Mozambique	17
1.4 The problem of malnutrition	19
1.4.1 <i>Malnutrition in Mozambique and Zambézia Province</i>	20
1.4.2 <i>TSNI project– Nutritional Status in the study area</i>	21
1.5 Nutrition education to improve health status of children under 5 in developing countries – does it work?	24
1.5.1 <i>Review of nutrition education studies in developing countries – their impact</i>	26
1.6 Lessons learned from theory based interventions in developed countries	32
1.7 Evidence-based effective nutrition interventions in developed countries	33
CHAPTER 2	35
FRAMEWORK MODEL.....	35
2.1 Persuasive communication in nutrition education.....	35
2.2 McGuire’s Communication/persuasion model	40
2.2.1 Systematic Literature Review	43
2.3 Literature review of nutrition education studies comparing both approaches and interventions	61
2.4 Aims and Research questions.....	67

CHAPTER 3	68
THE TSNI STUDY	68
3.1 TSNI project objectives, study setting and design	68
3.2 TSNI program activities and measurements collected.....	71
3.3 Characteristics of TSNI study population – the mothers	76
3.4 TSNI Socio-Demographic factors (or exposures)	77
3.5 Group class and individual class approaches in nutrition education – the TSNI case	79
3.5.1 <i>Input steps – the TSNI case</i>	79
3.5.2 <i>Channel – group and individual class processes and the TSNI case</i>	92
3.5.2.1 <i>Group class approach</i>	92
3.5.2.2 <i>Individual class approach</i>	94
3.6 Complex interventions	96
3.7 TSNI (published) results	99
3.7.1 <i>Impact from the nutrition education activities</i>	99
3.7.2 <i>Impact from the agriculture activities</i>	100
CHAPTER 4	102
METHODS AND TOOL DEVELOPMENT FOR THIS STUDY	102
4.1 Study population	105
4.1.1 <i>Sample size calculations</i>	106
4.2 Exposure	108
4.2.1 <i>Example of other studies' methods for measuring exposure</i>	108
4.2.2 <i>Methods for measuring exposure in this research study</i>	108
4.3 Attention	109
4.3.1 <i>What is attention?</i>	109
4.3.2 <i>Attention and learning/performance</i>	109
4.3.3 <i>Eye movements</i>	110
4.3.4 <i>Methods for measuring attention</i>	110
4.3.5 <i>Example of other studies' methods for measuring attention to health messages.</i>	111
4.3.6 <i>Methods used for measuring attention in this research</i>	113
4.4. Comprehension/Understanding.....	116
4.4.1 <i>The Nature of Understanding</i>	117

4.4.2 Example of other studies' methods for measuring comprehension to health messages.	119
4.4.3 Methods used to measure comprehension in the current research	121
4.4.3.1 Comprehension lexical meaning.....	121
4.4.3.2 Comprehension psychological-operational meaning.....	123
4.5 Skills acquisition	125
4.5.1 Example of other study's methods for measuring skills acquisition in health communications.	126
4.5.2 Methods used to measure skills acquisition in the current research	128
4.6 Memorisation/recall.....	131
4.6.1 What is Memory?	131
4.6.2 Measures of memory – recall:	132
4.6.3 Measuring recall to health messages:	133
4.6.4 Methods used to measure recall in the current research:	134
4.7 Attitude change	135
4.7.1 The nature of attitudes	136
4.7.2 Attitude formation and change	136
4.7.3 The attitude-behaviour relationship	137
4.7.4 Development of the attitude instrument in the current research study	138
4.7.4.1 Mode of administration	138
4.7.4.2 Devising the items	139
4.7.4.3 Scale	144
4.8. Data analyses	147
4.8.1 Measuring the distribution	147
4.8.2 Tests for the comparison of groups	148
4.8.3 Tests for correlations and associations	149
4.8.4 Missing values	151
4.8.5 Measure of persuasiveness	154
4.8.6 Software packages	155

CHAPTER 5.....	156
RESULTS	156
5.1 Exposure	156
5.1.1 <i>Characteristics of study subjects.....</i>	<i>156</i>
5.1.2 <i>Attendance rates to group and individual classes.....</i>	<i>157</i>
5.1.2.1 <i>Group class attendance</i>	<i>158</i>
5.1.2.2 <i>Individual class attendance</i>	<i>159</i>
5.1.3 <i>Exposure variables affecting attendance levels.....</i>	<i>160</i>
5.1.4 <i>Comparing attendance rates between approach-intervention groups.....</i>	<i>162</i>
5.2 Attention	162
5.2.1 <i>Exposure factors affecting attention</i>	<i>164</i>
5.2.2 <i>Comparing observed attention scores between intervention groups and between approach-intervention groups</i>	<i>167</i>
5.2.3 <i>Comparing results for attention-recall between intervention groups and between approach-intervention groups</i>	<i>170</i>
5.3 Understanding/Comprehension.....	173
5.3.1 <i>Comprehension- lexical meaning</i>	<i>173</i>
5.3.1.1 <i>Exposure variables affecting comprehension-lexical meaning.....</i>	<i>173</i>
5.3.1.2 <i>Comparing comprehension- lexical meaning between approach-intervention groups.....</i>	<i>174</i>
5.3.2 <i>Comprehension- operational and psychological meaning</i>	<i>178</i>
5.3.2.1 <i>Exposure variables affecting comprehension- psychological and - operational meaning.....</i>	<i>178</i>
5.3.2.2 <i>Comparing comprehension- psychological and -operational meaning scores between the intervention groups</i>	<i>180</i>
5.4 Skills acquisition	182
5.4.1 <i>Exposure variables affecting skill acquisition scores</i>	<i>183</i>
5.4.2 <i>Comparing skill acquisition scores between intervention groups</i>	<i>184</i>
5.5 Recall.....	186
5.5.1 <i>Exposure variables affecting recall scores</i>	<i>187</i>
5.5.2 <i>Comparing recall scores between approach/intervention group</i>	<i>187</i>
5.6 Attitudes.....	189
5.6.1 <i>Association between attitude sum scores and exposure factors.....</i>	<i>191</i>

5.6.2 Attitudes towards breastfeeding: comparison between the two intervention groups and with the control group.....	192
5.6.3 Attitudes towards complementary feeding: comparison between the two intervention groups and with the control group.....	193
5.6.4 Comparing summed scores for each section between the intervention groups and control group.....	195
5.6.5 Comparing section A to section B for each group.....	199
5.7 Associations between the output steps	199
5.8 Comparing persuasiveness between the two Intervention groups.....	205
CHAPTER 6	207
DISCUSSION AND CONCLUSIONS	207
6.1 McGuire's model and persuasive communication	208
6.1.1 Understanding the communication process.....	208
6.1.2 User-friendliness of McGuire's model.....	209
6.1.3 Delivery methods as input steps in the model	209
6.1.4 Exploring the model	211
6.1.5 Future use of McGuire's model for evaluation	213
6.1.6 Summary of findings for this section:.....	213
6.2 The developed tools.....	214
6.2.1 Exposure/attendance measurement.....	214
6.2.2 Attention measurement.....	216
6.2.3 Comprehension measurement	219
6.2.4 Recall and skills acquisition measurement	223
6.2.5 Attitude measurement.....	227
6.2.6 Measuring reliability and validity.....	229
6.2.7 Summary of findings for this section.....	230
6.3 Which intervention group is more effective?.....	231
6.3.1 Exposure/attendance rates	231
6.3.2 Block 2 output steps.....	233
6.3.3 Block 3 output steps.....	238
6.3.4 Attitudes towards breastfeeding and complementary feeding practices	241
6.3.5 Degree of persuasiveness	243
6.3.6 Comparing TSNi results between the two intervention groups	244
6.3.7 Which intervention group is better?	245

6.3.8 Summary of findings for this section.....	248
6.4 Overall main study design limitations	250
6.5 CONCLUSIONS	252
References	254
APPENDICES	267

LIST OF TABLES

Table 1.1	<i>Poverty and health indicators in Mozambique and Zambézia Province</i>	21
Table 1.2	<i>Nutritional status of under five reference children at baseline</i>	23
Table 2.1	<i>Description of studies that have used McGuire's communication/persuasion model to evaluate a health education program</i>	48
Table 2.2	<i>Description of studies comparing a group versus individual nutrition education approaches and the two intervention groups</i>	64
Table 3.1	<i>Topics given in the group classes</i>	81
Table 4.1	<i>Summary of methods and sample size for each output step</i>	146
Table 4.2	<i>Tests for association between output steps and exposure variables</i>	152
Table 5.1	<i>Characteristics of TSNi study population by intervention group (n=560)</i>	157
Table 5.2	<i>Descriptive statistics for attendance by approach-intervention group</i>	158
Table 5.3	<i>Associations between Socio-Demographic variables and attendance to group and individual classes</i>	161
Table 5.4	<i>Analysis of Variance of attendance levels</i>	162
Table 5.5	<i>Characteristics (mean and SD) of variables (i.e. exposure factors) measured during attention measurements</i>	163
Table 5.6	<i>Attention levels measured with both methods in 10 cases on two occasions</i>	164
Table 5.7	<i>Association between exposure variables and attention levels</i>	166
Table 5.8	<i>Descriptive statistics for attention observation by approach-intervention group</i>	167
Table 5.9	<i>Analysis of Variance for attention scores</i>	168
Table 5.10	<i>Adjusting the association between approach-intervention and attention scores for class duration and disturbance levels (using multiple regression)</i>	169
Table 5.11	<i>Adjusting the association between approach-intervention and attention-recall for class duration and disturbance levels (using logistic regression)</i>	171

Table 5.12	<i>Association between exposure variables and understanding</i>	173
Table 5.13	<i>Adjusting the association between approach-intervention groups and comprehension scores for class duration and disturbance levels (using logistic regression)</i>	176
Table 5.14	<i>Descriptive statistics for exposure variables and comparison between intervention groups</i>	179
Table 5.15	<i>Association between comprehension levels and socio-demographic characteristics and with other steps</i>	180
Table 5.16	<i>Descriptive statistics for comprehension</i>	180
Table 5.17	<i>Adjusting the association between intervention groups and comprehension scores for confounding (using regression)</i>	182
Table 5.18	<i>Recipe type and ingredients demonstrated per locality</i>	183
Table 5.19	<i>Association between skills acquisition and exposure variables</i>	184
Table 5.20	<i>Descriptive statistics for Skills acquisition scores</i>	184
Table 5.21	<i>Adjusting the association between intervention group and skills acquisition scores for possible confounders (using regression)</i>	186
Table 5.22	<i>Association between recall and exposure variables</i>	187
Table 5.23	<i>Descriptive statistics for recall scores</i>	187
Table 5.24	<i>Adjusting the association between intervention group and recall scores for confounders (using regression)</i>	189
Table 5.25	<i>Analysis of Variance for socio-demographic variables</i>	190
Table 5.26	<i>Association between attitude scores and socio demographic variables by intervention groups (using Pearson's r and Independent T-test (CI))</i>	191
Table 5.27	<i>Comparing proportion of answers to attitude questions regarding breast feeding practices between the two intervention groups (intervention I n=139, Intervention II n=142 and control n=130)</i>	194
Table 5.28	<i>Comparing proportion of answers to attitude questions regarding complementary feeding practices between the two intervention groups (intervention I n=139, Intervention II n=142 and control n=130)</i>	194
Table 5.29	<i>Descriptive statistics for Section A's scores for each treatment group</i>	195

Table 5.30	<i>Analysis of Variance for the sum of scores - Breastfeeding attitudes</i>	196
Table 5.31	<i>Descriptive statistics for Section B's scores by treatment group</i>	196
Table 5.32	<i>Adjusting the association between attitude scores and treatment group (multiple regression)</i>	198
Table 5.33	<i>Adjusting the association between intervention groups and attitude scores (multiple regression)</i>	198
Table 5.34	<i>Associations between output steps for each approach-intervention group</i>	200
Table 5.35	<i>Weights allocated to each output step</i>	205
Table 5.36	<i>Descriptive statistics of the weighted scores for each output step by intervention group</i>	206

LIST OF FIGURES

Figure 1.1	<i>Map of Mozambique</i>	18
Figure 1.2	<i>Causal Framework outlining the determinants of malnutrition</i>	20
Figure 2.1	<i>The nutrition communication model</i>	38
Figure 2.2	<i>Selection of studies</i>	47
Figure 3.1	<i>Study areas</i>	71
Figure 3.2	<i>TSNI conceptual framework</i>	72
Figure 3.3	<i>Timing of TSNI research and extension activities: 2003-2005</i>	74
Figure 3.4	<i>Recipe demonstrations with the women's groups</i>	83
Figure 3.5	<i>Puppet show</i>	84
Figure 3.6	<i>Hand-outs of the 6 recommendations given to mothers</i>	86
Figure 3.7	<i>Group class</i>	94
Figure 3.8	<i>Individual class session</i>	96
Figure 3.9	<i>MRC Framework model for the evaluation of complex interventions</i>	98
Figure 5.1	<i>Attendance level at each class (topic) by intervention group</i>	158
Figure 5.2	<i>Total number of group classes attended by approach-intervention group</i>	159
Figure 5.3	<i>Correlation between Intervention II attendance to group classes and individual classes</i>	160

Figure 5.4	<i>Observed attention scores by approach-intervention group</i>	168
Figure 5.5	<i>Attention-recall scores by approach-intervention group</i>	170
Figure 5.6	<i>Proportion of question type administered to mothers in each of the approach-intervention groups</i>	174
Figure 5.7	<i>Comprehension (Lexical meaning) responses by approach-intervention groups</i>	175
Figure 5.8	<i>Comparing Comprehension (Psychological and Operational meaning) scores between intervention I and II</i>	181
Figure 5.9	<i>Comparing skills acquisition scores between intervention groups</i>	185
Figure 5.10	<i>Comparing recall scores between intervention groups</i>	188
Figure 5.11	<i>Comparing distribution of sum of breastfeeding- attitude scores among groups</i>	195
Figure 5.12	<i>Comparing distribution of sum of complementary feeding-attitude scores among treatment groups</i>	197

LIST OF APPENDICES

Appendix 4.1	<i>Tool used to measure attention observations, attention recall and comprehension-lexical meaning in group classes – an example</i>	267
Appendix 4.2	<i>Tool used to measure attention observations, attention recall and comprehension-lexical meaning at individual classes – an example</i>	269
Appendix 4.3	<i>Tool used to measure comprehension-psychological meaning, skills acquisition and recall – and example</i>	271
Appendix 4.4	<i>Attitude tool – an example</i>	274
Appendix 4.5	<i>Faces used as the Likert scale</i>	277
Appendix 4.6	<i>Example of vignettes and corresponding story used in the attitude tool</i>	278
Appendix 4.7	<i>Example of answers given when attitudes to complementary feeding items were negative</i>	280

CHAPTER 1 — INTRODUCTION

MALNUTRITION AND NUTRITION EDUCATION: BACKGROUND OF STUDY

1.1 Overview and purpose of the study

“Malnutrition has been responsible, directly or indirectly, for 60% of the 10.9 million deaths annually among children under five.” [1]

For decades there has been intense interest in identifying effective strategies for reducing malnutrition in infants and young children in developing countries and nutrition education has been one of the candidates. However, it is in developing countries that the challenge of providing nutrition education is often the greatest.

In a country like Mozambique with scarce resources and extremely poor under five nutritional status, it is essential to evaluate the most common delivery methods to nutrition education or a combination of approaches in order to choose the most effective.

At the same time, due to the uncertain contribution of nutrition education to behaviour change, it is imperative that we understand why a health message leads (or does not lead) to a person changing her/his behaviour. Evaluations of nutrition education programmes mainly focus on outcomes such as behaviour change and change in nutritional status but rarely take into account earlier stages of the communication process. Among the many models that guide current research and development in health communications, McGuire’s communication/persuasion model is one of the most practical for designing strategies for monitoring and evaluation. The model is based on communication “inputs” (which make up the message side of the communication) and “outputs” (which make up the receiver’s side of the communication). Of more direct relevance to monitoring and evaluation are the “output-steps”: exposure, attention, understanding, skills acquisition, memorisation/recall, attitudes and behaviour change.

This study has two primary objectives: carrying out a process evaluation of a nutrition education program in Mozambique, which will allow for the understanding of the

communication process. McGuire's model is in this case used as the conceptual framework, contributing to an innovative way of evaluating nutrition education programs in developing countries. Secondly, along with the first objective, this study makes use of the model to compare two nutrition education delivery methods: a group class only intervention with a group class plus individual class intervention, hence looking at the added value of the individual class approach.

This study also contributes by developing and empirically testing tools for the measurement of McGuire's output steps in a rural setting of a developing country.

Results of this study might possibly illuminate a black box for nutrition education specialists by providing an improved understanding of the multi-step communication process, and also by contributing with information on the effectiveness of an approach or combination of approaches. Moreover, an improved understanding of these aspects of nutrition education programs could assist policy makers to make appropriate decisions regarding resource allocation for nutrition education programs in Mozambique and other resource-poor settings.

The current study was linked to the nutrition education program of a much bigger study, the TSNI study. Measurements developed only for this PhD study were collected at the same time and as part of the nutrition education classes of the TSNI study. The current study looks at the process evaluation of the communication section of the TSNI nutrition education program, but does not attempt in any way to link its results to the TSNI main results. Results from one study are to be seen separately from the results of the other study. One reason is that the sample size measured in the current study is a much smaller sub set of the TSNI study. The only outcome which would have been of the outmost relevancy for the current study would have been behaviour change, however, the TSNI results for this outcome are incomplete.

1.2 Contents

This thesis is divided into 6 chapters. Chapter 1 provides the background information for the reader. After stating the problems of child malnutrition in Mozambique, it presents a review of the literature on nutrition education in developing countries (DCs) including the evidence of its benefits. Chapter two describes the conceptual framework for the study and presents a systematic review of studies that have used McGuire's model for a similar purpose. The second part of the chapter presents a review of health education studies that have compared the group and individual nutrition education approaches. It finishes by providing the aims, objectives and hypothesis of the thesis. Chapter three describes the background literature for existing tools used to measure each of McGuire's output steps and the methods developed to measure them in this study. Chapter four describes the study setting, study population, sample size calculations, tests used for data analysis, and also describes the two nutrition education approaches measured in this study. Chapter 5 presents the results for each output step. Chapter six concludes the thesis with a general discussion of the main findings, lessons learned, recommendations and conclusions.

1.3 Background of Mozambique

This study was carried out in Zambezia Province, Mozambique, represented in figure 1.1 (Zambézia Province is outlined by the bold black lines).

This map illustrates the geographical context of Mozambique, showing its international borders and internal administrative divisions. Key features include:

- Neighboring Countries:** Zambia to the north, Tanzania to the northeast, Zimbabwe to the west, South Africa to the southwest, Swaziland to the south, and Botswana to the west.
- Major Cities and Towns:** Harare (Zimbabwe), Lilongwe (Malawi), Maputo (South Africa), Beira, Vila Rica, Inhassoro, Vilanculos, Nova Mambone, Espungabera, Dombe, Chicualacuala, Massingir, Inharrime, Xai-Xai, Manhica, and Maputo.
- Rivers and Water Bodies:** Zambezi R., Save R., Limpopo R., Orange R., Lake Malawi, Lake Kariba, Lake Bangweulu, and the Indian Ocean.
- Geographical Features:** The map highlights the Mozambique Channel and the Mozambique Strait.
- Scale and Orientation:** A scale bar indicates distances up to 200 km. A compass rose shows the cardinal directions.
- Map Source:** ©1997 MAGELLAN GeographixSM (805) 685-3100 www.maps.com

Since then the country has enjoyed political stability and significant economic growth. Foreign investment and tourism are now being promoted in order to build stability. However, the country remains dependent upon foreign assistance for much of its annual budget, and around 69% of the population remains below the poverty line.

It is still one of the world's poorest countries, with a per capita gross domestic product (GDP) of US\$210 in 2000. It has a low population density (about 20/km²), a official population estimate of 18,500,000 for 2003, a mainly rural population (69.5% in 2003) and a predominantly agricultural economy (30 percent of the GDP and 80 percent of exports) that is particularly vulnerable to the recurring natural disasters (cyclones, floods and drought) to which the country is prone. Sixty-four percent of all Mozambicans live in food-insecure households. Mozambique's human development indicators are among the worst in the world. [2-4]

1.4 The problem of malnutrition

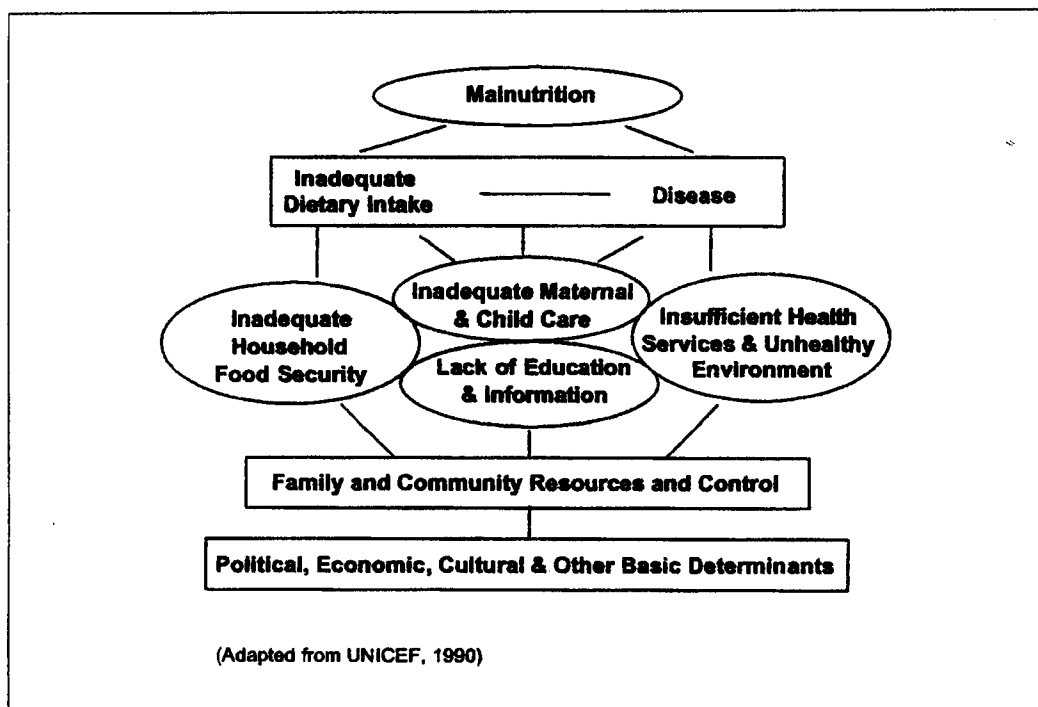
According to UNICEF and others, all people have the right to be free from hunger and malnutrition. Despite this basic human right, over 200 million children in developing countries under the age of five are malnourished. Malnutrition contributes to more than half of the nearly 12 million under-five deaths in developing countries each year.[5]

Malnutrition can derive from micro- and/or macro-nutrient deficiencies. Micronutrients are substances that the human body cannot make itself, such as vitamins and minerals, which are needed in very small amounts (micrograms or milligrams per day), as distinct from fats, carbohydrates, and proteins which are macronutrients, since they are needed in considerably greater amounts.[6]

Among the most common and significant micronutrient deficiencies affecting children in developing countries are iron, iodine and vitamin A deficiencies (VAD). Of these, VAD is the single most important cause of childhood blindness in Developing Countries. It also contributes significantly, even at sub clinical levels to morbidity and mortality from common childhood infections. VAD is the result of two primary factors: persistent inadequate intake of Vitamin A that is frequently exacerbated by other dietary circumstances, and a high frequency of infections. An estimated 2.8 million pre-school age children are at risk of blindness from VAD, and the health and survival of 251 million others are seriously compromised.[7, 8]

There has been substantial international experience documenting the determinants of child malnutrition. These causal factors have been integrated into the often-cited UNICEF Conceptual Framework published in 1990 (figure 1.2).[9]

Figure 1.2. Causal Framework outlining the determinants of malnutrition.[9]



1.4.1 Malnutrition in Mozambique and Zambézia Province

Apart from understanding the consequences of the poverty situation on the health of the population in Mozambique, it is also important to show the same figures for Zambézia Province since the study described in this thesis was carried out in three districts of this Province.

According to WHO figures (2000), Mozambique is within the highest mortality stratum of African countries.[10] Zambézia is one of the poorest provinces in Mozambique, it is also the second most populated in the country.[11] As can be seen from Table 1.1, child and infant mortality rates and malnutrition rates are very high in Mozambique. Malnutrition rates (stunting, wasting and underweight) in children under five years of age were even higher for Zambézia Province.[12]

High levels of vitamin A (VAD) and iron deficiencies also characterize the nutrition situation in Mozambique. A nationally representative survey for Mozambique was carried out between December 2001 and February 2002 by the Mozambican Ministry of Health to

determine the prevalence of Vitamin A deficiency (measured by assessing serum retinol status), anaemia, and of malaria in children (6-59 months of age) and their respective mothers.[13] Subsequent results are shown in Table 1.1.

Table 1.1 Poverty and health indicators in Mozambique and Zambézia Province

General	National	Zambézia Province
Total Population (2003) [2]	≈8.5 million	≈3,4 million
Under five mortality rate (2003) [12]	153 ⁴	123 ⁵
Infant mortality rate (2003) [12]	101 ⁴	89 ⁵
Under 5 stunting ¹ rate (<-2 z-scores) [12]	36.7%	47.3%
Under 5 wasting ² rate (<-2 z-scores) [12]	4%	5.2%
Under 5 underweight rate ³ (<-2 z-scores) [12]	23.7%	26.9%
Maternal malnutrition (BMI= <18.5) [12]	8.2%	11%
Severe VAD in under 5 (<0.35 µmol/L) [13]	13.8	N.A
Moderate VAD in under 5 (<0.70 µmol/L) [13]	55%	N.A
Light anaemia in under 5 (>10g/dl -- <11g/dl) [13]	23.2%	N.A
Moderate anaemia in under 5 (10g/dl – 7g/dl) [13]	44%	N.A
Severe anaemia in under 5 (< 7g/dl) [13]	7.2%	N.A

¹ Stunting (HAZ) is used in children whose height for a particular age is more than two standard deviations below the median value of the reference population. [14, 15]

² Wasting (WHZ) is used in children whose weight for a particular height is more than two standard deviations below the median value of the reference population. [14, 15]

³ Low weight for age (WAZ) is the third indicator of malnutrition and is the most commonly used in the developing world, because it does not involve the measurement of height, which requires more extensive training to do accurately under field conditions. [14, 15] ⁴ Death rates for the period of 1998 -2003; ⁵ Death rates for the period of 1993-2003;

1.4.2 TSNI project– Nutritional Status in the study area

The current research study (i.e. this thesis) was linked to a large randomized controlled trial entitled: “Towards Sustainable Nutrition Improvement (TSNI): Addressing Macro- and Micro-nutrient Malnutrition Through New Cultivars and New Behaviours”, based in Zambézia Province (figure 1.1), with a duration of approximately 2.5 years (October 2002-March 2005). The project integrated nutritional, agronomic, and socio-economic components to achieve the project objective: to determine whether a food-based

intervention strategy can lead to sustainable, year-round intake of vitamin-A rich foods, reduced fluctuations in seasonal household calorie supply, and an overall improvement of diet diversity, nutritional status and diet quality in a cost-effective manner, particularly among children under five years of age. Emphasis was placed on addressing two major nutritional problems: vitamin A deficiency and inadequate caloric intake. To achieve this, the project introduced beta-carotene rich varieties of sweet potato, utilizing these new cultivars as a low-cost, effective entry point for improving the kinds of weaning foods given to young children and increasing the frequency of intake of essential micro-nutrients and calories.[16]

Nutritional status of under five reference children in the study was measured at baseline by collecting anthropometric data and blood to measure serum retinol and haemoglobin levels. As can be seen from table 1.2, the study was working in sites with very high prevalence of malnutrition levels in under five. In terms of the proportion of the total sample, 23.3 % had a low Hb status, 10.6% were seriously deficient in serum retinol (<0.35 $\mu\text{mol/L}$), and 71.2% were deficient in serum retinol (<0.70 $\mu\text{mol/L}$), 54.4% were chronically (HAZ) malnourished, 8.2% were acutely malnourished (WHZ), 49.8% had low weight for age and 30.3% had low arm circumference. [16]

Further results for the TSNI study can be viewed at:

http://www.aec.msu.edu/fs2/tsni/TSNI_FINDINGS1_TEXT_DONOR_REPORT.pdf

Table 1.2 Nutritional status of under five reference children at baseline[16]

NUTRITIONAL CHARACTERISTICS	INTERVENTION children (N=498)	
	MEAN	STD DEV.
HEMOGLOBIN (g/l)	8.4	1.7
SERUM RETINOL (umol/L)	0.611	0.223
CHRONICALLY MALNOURISHED (HAZ)	2.06	1.17
ACUTELY MALNOURISHED (WHZ)	0.72	0.89
LOW WEIGHT-FOR-AGE (<-2 Z-SCORE)	1.89	1.07
LOW ARM CIRCUMFERENCE FOR AGE ¹ (<-2 Z-SCORES)	13.76	1.41

¹ Mid - Upper Arm Circumference (MUAC), is relatively easy to measure and a good predictor of immediate risk of death. It is used for rapid screening of acute malnutrition from the 6-59 month age range. [14, 15]

Apart from introducing beta-carotene rich varieties of sweet potato, the TSNI study also carried out a nutrition education program. As a research component, the study aimed to compare the effectiveness of two nutrition education interventions: group class approach (intervention I) versus a group class and individual counselling approach combination (intervention II). Effectiveness of the nutrition component in the TSNI study was measured mainly by looking at nutritional status of under five children and their nutrient intake, and secondly, at reported dietary practices.

It would have made more sense for the TSNI study to have a third arm which would be an intervention consisting of the individual counselling approach only. However, an intervention with three arms as such would be much more costly. Another possibility would have been to compare a group class only approach with an individual counselling only approach, which makes more sense since the latter is less costly and less time consuming than having a group class plus individual counselling arm. However, the TSNI wanted to specifically look at the added advantage of receiving the individual counselling at home, hence the design that it used. It is not possible to discuss at this point about which of the two interventions is more cost effective since the TSNI study has not yet published any results comparing results from both intervention groups. However, the cost of implementing intervention II was much higher than intervention I due to the individual counselling approach.

From this point forward in this thesis there will be a clear distinction between the terms intervention groups (referring to intervention I and intervention II) and approaches (referring to group class and individual class approaches). The term approach-intervention group will also be used from the results section onwards to refer to an approach within an intervention group, for example, group class-intervention I or group class-intervention II. Moreover, the term “individual counselling” as used in the TSNi study will be referred to as “individual class” in the current study as it was more of a class (i.e. a teaching) experience than a “consultation/counselling” which would suggest discussion and diagnosis.

My role in the TSNi study was as the person responsible for developing and monitoring the nutrition education program of the TSNi study. At the same time I carried out my own PhD research study which was linked to the TSNi study. All tools (with exception of exposure) described in chapter 4 of this thesis were developed by myself specifically for my PhD research study and were not part of the TSNi study.

1.5 Nutrition education to improve health status of children under 5 in developing countries – does it work?

As discussed in previous sections, malnutrition in developing countries is widespread, especially in very young (i.e. under five) children. Improving the nutrition status of these children is therefore a humanitarian and economic imperative. Nutrition education could be a promising strategy to combat malnutrition. Nutrition education can be used as a tool to provide people in rural and urban areas with adequate information, skills and motivation to procure and to consume appropriate diets. Such education should cover improvement of family food supplies and more efficient utilization of available food and economic resources to provide nutritious diets and better care for the most vulnerable groups.[17]

As defined by Contento *et al* (1995): “nutrition education is any set of learning experiences designed to facilitate the voluntary adoption of eating and other nutrition-related behaviours conducive to health and well-being”[18]. This suggests that behavioural change is the ultimate criterion for the effectiveness of nutrition education. A common example of a desired behaviour change that would be the focus of nutrition education in developed countries would be the decrease in the consumption of high-fat and high-sugar foods. Other examples, appropriate for developing countries, include increasing the

consumption of vitamin A-rich fruits, dark green leaves (rich in iron and folic acid), and good-(quality) sources of protein and iodine.

In many occasions, nutrition education might be replaced by the terms health promotion or health education. It is well established in the general health promotion community that there are distinct differences between the latter two terms.[19, 20] Health education “is an activity that seeks to inform the individual on the nature and causes of health/illness and that individual's personal level of risk associated with their lifestyle-related behaviour. Health education seeks to motivate the individual to accept a process of behavioural-change through directly influencing their value, belief and attitude systems, where it is deemed that the individual is particularly at risk or has already been affected by illness/disease or disability.” [21] Health education focus more at individuals and changing their behaviour, while health promotion tends to focus more on changing organizations and policies. [21] [19, 20] Health promotion “is the process by which the ecologically-driven socio-political-economic determinants of health are addressed as they impact on individuals and the communities within which they interact. This serves to counter social inaction and social division/ inequality. It is an inherently political process that draws on health policy as a basis for social action that leads to community coalitions through shared radical consciousness. Health promotion seeks to radically transform and empower communities through involving them in activities that influence their public health – particularly via agenda setting, political lobbying and advocacy, critical consciousness-raising and social education programmes. Health promotion looks to develop and reform social structures through developing participation between representative stakeholders in different sectors and agencies.”[21] As can be seen from the two definitions, health promotion is more directed to ecologically and environmentally-driven actions and community-based empowerment, while health education has a more individualistic orientation. The TSNI study for example, had a component of both health education and health promotion. Because the study tried to deal with social norms through social marketing, hence not only tried to changed individual practices but also the group and community practices, it looked at the cultural and ecological context. But it also worked with individuals with the idea of empowering them to change their behaviour, and hence, the health education influence.

An important public health question is whether educational programmes to improve complementary feeding and breastfeeding practices will reduce malnutrition in under five children.

A review by Ashworth and Feachem (1985)[22] “of data from 12 developing countries suggests that it is possible, even in poor communities, to improve substantially the nutritional status of infants and young children by weaning education.”[22] They also mentioned that “in the few instances where nutrition education has been evaluated, the criteria chosen have usually been either an improvement in nutrition knowledge or a change in attitude.” [22]

A report by FAO (1991)[23] which reviewed the status of nutrition education programs, argued that many nutrition education efforts implemented during the 1970s and early 1980s reported few changes in nutrition-related behaviour or nutritional status. The report concluded that programme failure could be attributed largely to the ineffective communication methods used and the inappropriate content of the messages, which often ignored the specific cultural context of the community.

Since then there have not been any reviews on the current status of nutrition education programmes in developing countries. Nonetheless, Sheikholeslam *et al* (2004)[24] have discussed that Communication for Behaviour Change programmes on a large scale, especially those aiming at mothers and others who care for children, have shown they were effective in improving the nutritional status of children.

The following is a review of studies specifically on breastfeeding and complementary feeding nutrition education programmes in developing and middle-income countries. It is an attempt to inform on the present status of this type of studies in terms of their effectiveness and other factors related to it.

1.5.1 Review of nutrition education studies in developing countries – their impact

This literature review reports the findings of studies that have used nutrition education to improve directly or indirectly the health/nutrition status of children under 5 in a developing country, or middle-income country (only studies targeting subjects of low socio-economic status were selected). The main objective is to determine the use of approaches and frameworks in nutrition education, topics that shall be discussed in the following chapters.

The TSNI's nutrition education programme specifically targeted breast feeding and complementary feeding practices. For this reason, this review will only look at nutrition education controlled studies promoting breast-feeding and complementary feeding/weaning practices. This review will try to report the following:

1. If any studies have explicitly indicated the use of any models or theories or elements of models/theories to support the nutrition education program, especially in explaining what factors (for example measures of potential psychosocial mediators or correlates of behaviour) made it work (i.e. evaluating its success). This objective follows the reasoning by Smith and Lopez (1991)[25]: "the current focus in nutrition education research is not only on whether nutrition education is effective, but on what makes it work" and "basing programs on theory might help clarify the factors that mediate their effectiveness".
2. Determine the types of nutrition education approaches used in the reviewed studies.
3. List the outcomes that have been commonly evaluated
4. The success of these studies

Keywords, titles and abstracts of PubMed were searched using the various search terms. Additionally, the reference sections of the relevant articles were hand-checked to identify any additional studies and, related articles retrieved from PubMed were looked through for any relevant studies.

Search terms:

1. Breast feeding promotion studies

- *("Breast Feeding"[MeSH]) AND ("Health Promotion"[MeSH] OR "Health Education"[MeSH]) AND ("Developing Countries"[MeSH])*
- *Breast feeding promotion AND Developing Countries*

2. Complementary feeding promotion studies

- *"Health Education"[MeSH] AND "Developing Countries"[MeSH] AND "Nutrition Processes"[MeSH]*
- *Nutrition education AND developing countries AND children*

Because the above two search terms produced no results or irrelevant studies, related articles to a relevant study[26] were instead looked at in PubMed.

Results:

A. Breastfeeding promotion studies

A total of 15 breastfeeding promotion studies[27-41] were identified. Only those studies that compared the outcomes between intervention group and control group were included in this review. Within group comparisons were not included.

1. The use of a model/theory

No studies reported the use of a model or theory as a framework to evaluate the intervention, especially looking at factors underlying the success of their programmes, answering the question *what made it work?*

2. Nutrition education approaches used

Regarding the channels of communication used, the majority of studies reported using one-to-one education, be it in a health clinic, or at home or as peer counselling. Only 5 studies used group education. Two studies apart from using one-to-one or/and group class approaches also reported using posters in health clinics and printed health information. The latter is more difficult in developing countries where literacy rates may be low. Because of the criteria that studies had to include a control group, the use of channels in these studies was restricted to face-to-face communications, and hence, this review left out mass communications which are also widely used. A review by Green's (1999)[42] looking at key breastfeeding behaviours only described 3 channels of communication: one-to-one, women's groups (or group classes), and mass media. Regarding one-to-one communication she concludes that "hospital counselling (especially individual counselling) can have a dramatic and positive effect on exclusive breastfeeding rates. Once women are discharged from the hospital, home visits, including peer counselling, appear to play a critical role in supporting women's breastfeeding behaviours and are influential in extending the period of exclusive breastfeeding." [42] Regarding group classes, she also

mentions that little research has been carried out on the breastfeeding impact of support groups. Regarding mass media she concludes that “there are few definitive studies on the impact of mass media on breastfeeding behaviours. Of the studies that were sufficiently rigorous to draw conclusions about impact, none of them singled out mass media relative to other interventions.” [42]

3. Outcomes measured

Breastfeeding practices was the outcome most commonly measured. Four studies reported the measurement of nutritional status and only 1 study reported the measurement of knowledge and attitudes. The latter was further supported by another search in PubMed using the search terms: *breastfeeding AND knowledge AND (health promotion OR health education) AND developing countries; "Breast Feeding"[MeSH] AND ("Health Promotion"[MeSH] OR "Health Education"[MeSH]) AND "Health Knowledge, Attitudes, Practice"[MeSH] AND "Developing Countries"[MeSH]; breastfeeding AND attitudes AND (health promotion OR health education) AND developing countries*. Of the retrieved studies only 2 measured attitudes and knowledge.

The above findings was also supported by the fact that a review by Green[42] on breastfeeding promotion studies found that studies tended to measure the impact on breastfeeding behaviours/practices and only a few studies measured nutritional status of children – “the ultimate test of effective impact.” [42]

4. Success of studies

The majority of these studies reported to have succeeded in almost all outcomes measured.

This is supported by Dewey [43]: “Nutrition-education trials that included an emphasis on breastfeeding, not just improved complementary foods, have generally been positive.” [43]

Even though this might lead to the conclusion that breastfeeding programmes have been successful, it is important to mention that this review did not look at the methodological limitations and many of the reviewed studies had flawed designs making it impossible to sort out what works and what does not, inadequate controls, failure to account for

confounding variables, and small sample sizes. Moreover, this review does not account for publication bias.

B. Complementary feeding promotion studies

A total of 9 complementary feeding promotion studies[26, 44-51] were identified. Only those studies that compared the outcomes between intervention group and control group were included in this review. Within group comparisons were not included. Studies that also included other types of intervention, for example, supplementation, were not included as it would make it difficult to discern if the benefit came from the nutrition education intervention or from the supplementation.

1. The use of a model/theory

Only one study⁴⁴ reported using a theory (Hubley's (1993) behavioural change model) to support their nutrition education programme in order to address change in nutrition behaviour. They measured beliefs, attitudes, subjective norms and enabling factors in order to explain behaviour change.

2. Nutrition education approaches used

Regarding the channels of communication used, once more, the majority of studies reported using one-to-one education, be it in a health clinic, or at home or as peer counselling. They have also reported using the group class approach.

3. Outcomes measured

The outcomes measured can be categorised into feeding practices, dietary intakes, and growth. Some studies also measured nutritional knowledge and attitude. Compared with the breastfeeding promotion studies, a bigger proportion of the complementary feeding studies measured growth.

4. Success of studies

The majority of studies reported to have succeeded in almost all outcomes measured. However, it is also possible that those studies that did not succeed were also not published (i.e. leading to publication bias).

This review shows that nutrition education studies looking specifically at breastfeeding and complementary feeding promotion in developing- and middle-income countries have not

been using models/theories as a framework to support their studies and therefore would not be able to explain their successes. As mentioned by Contento[52]: "It is also important to include studies that provide information on a variety of potential mediating factors that might contribute to the achievement of behavioural change goals in interventions. An understanding of the impact of interventions on mediating variables would help nutrition educators identify those factors that contribute to effectiveness". Maybe because of the lack of use of models/theories, the authors of these studies have neglected the measurement of certain outcomes like for example knowledge and attitudes.

Moreover, even though the review shows that the majority of studies have succeeded in their aimed outcomes these studies measure practices and not nutritional status which is a much better indicator of child health status. Moreover, practices have been measured by recall questionnaires and only a few studies used observations.

This review has indicated that the two most used delivery methods in breastfeeding and complementary feeding promotion studies in developing countries have been: face-to-face communications, and mass communication. The majority of studies have used face-to-face communications, specifically the one-to-one approach and the group class approach. None of the studies, however, compared these two approaches.

This section provides the background for the next chapter. The reader goes to the next chapter knowing that there is a need to use models/theories as a framework to support the intervention and that preferably, these models/theories should help us measure other factors/outcomes that could explain behaviour change. Attitudes and knowledge are two of these factors, but are not the only ones. This section also informed us about the education approaches that have commonly been used in breastfeeding and complementary feeding promotion studies in DCs and that there is currently no information available on which approach or what combination of approaches is most effective in a DC setting.

1.6 Lessons learned from theory based interventions in developed countries

Due to an absence of theory based interventions in developing countries, it is perhaps important to describe here some of those lessons learned from theory based interventions in nutrition education from developed countries which could be adapted to a developing country context.

There are 3 main groups of theories/models that have been used to explain behaviour change[53]:

- Those that focus on individual characteristics such as the Health Belief Model, the theories of Reasoned action and Planned Behaviour, the Transtheoretical model and the Social learning theory;
- Those that explain change in communities and communal action for health such as the diffusion of innovation theory;
- Those that guide the use of communication strategies for change to promote health, such as social marketing and McGuire's communication for behaviour-change.

Some mediating factors that can be selected from some of the above theories and which have shown to be relevant when designing interventions are [54]:

- Motivators and reinforcers that have personal meaning;
- Appropriate use of rewards;
- Expectations
- Knowledge
- Beliefs
- Self-efficacy
- Attitudes
- Behaviour intent
- Perceived threat and severity of disease
- Tailoring of messages
- Social support or community support

It has been shown for example, that communication efforts are more credible if they are culturally and personally relevant, hence, the need to tailor messages to a specific audience. [55] Moreover, studies on communication for prevention of HIV/AIDS find it

important to understand initial knowledge, attitudes, beliefs, perceived threat and severity, as well as expectations before designing messages.[56] Individuals, especially in DCs often bring with them a strong sense of identification with and attachment of individuals to their families and communities. The design of health messages must therefore acknowledge this.[57] Studies have also shown that subjects were able to change behaviours when self-efficacy (i.e. the subject believing that he/she is capable of taking the action) was a component of the intervention.[58]

1.7 Evidence-based effective nutrition interventions in developed countries

A report written by the Agency for Healthcare Research and Quality (AHRQ) [59] described what is known about the efficacy and effectiveness of behavioural interventions in promoting dietary change in cancer risk reduction. The report describes the results of a systematic review looking at the impact of behavioural intervention studies on the intake of fat and fruits and vegetables, which are relevant to cancer risk. Effectiveness and efficacy of the studies included in the systematic review was measured by analysing three factors: if there is evidence that maybe one type of intervention or maybe a combination of interventions is better than another (i.e. individual dietary counselling, group nutrition classes, social support groups); if interventions are more effective in certain subgroups of the population (i.e. children, youth, or old, or maybe people at risk, or even of different ethnic backgrounds); and, if conclusions can be reached about the cost-effectiveness of these types of interventions. [59]

Regarding those studies trying to increase the intake of fruits and vegetables, the main findings were:

Interventions were more successful at increasing fruit intake among children and vegetable among adults. In addition, interventions conducted among higher disease-risk populations were consistently more likely to report statistically significant increases in fruit and vegetable intake than were studies in general populations. [59]

Also studies employing a theoretical basis were more likely to report statistically significant increases in fruits and vegetable intakes than studies that did not use theory-based

interventions. Within theory components, the social support components was associated with more favourable increases in fruit and vegetable intake. [59]

Another important finding is that those studies that used interactive activities involving food were more likely to report statistically significant increases in fruit and vegetable intake. [59]

The authors did not have a large enough pool of articles to explore characteristics such as intervention intensity, setting, mode of delivery, use of individual tailoring, or culturally or ethnically specific interventions. [59]

Regarding those studies trying to increase the intake of fat, the main findings were: Interventions conducted among higher disease-risk populations were consistently more likely to report statistically significant increases in saturated fat intake compared to the general population, and interventions conducted among children appeared to be more successful at reducing intake of total fat and less successful at reducing intake of saturated fat than interventions conducted among adults. [59]

Unlike with interventions targeting an increase in fruit and vegetable intake, interventions targeting an increase in fat intake and which employed a theoretical framework were not consistently more likely to report significant effects. [59]

The use of social support, involving families and delivering the information in small groups, and goal setting appeared particularly effective at reducing intake of dietary fat. [59]

Although studies that used interactive food-related activities were more likely to report significant decreases in fat intake the magnitude of the decrease was not higher than among the studies that did not incorporate these special features. [59]

Overall, looking at the above, several dietary intervention components appear to be promising in modifying dietary change. These factors include: [59]

- Social support
- Goal setting
- Small groups
- Food-related activities.
- The incorporation of family components.

CHAPTER 2

FRAMEWORK MODEL

The first part of this chapter explains, describes and justifies the model used as the framework for the current research study. It also reports the findings of a systematic review of other studies that have also used the same model within health education/promotion.

As was seen in the previous chapter, the one-to-one approach and the group class approach are the two most used delivery methods in breastfeeding and complementary feeding promotion studies in developing countries, however, none of these studies have compared them. Hence, in an attempt to partially fill this gap, the second part of this chapter reports the findings of a review on general nutrition education studies that have compared a group class approach with an individual counselling approach, or a combination of these two approaches.

2.1 Persuasive communication in nutrition education

The focus now in nutrition education research has evolved from looking solely at the effectiveness of programs to looking at the reasons for why it did (or not) work.

As could be seen from the review in the previous chapter on nutrition education studies in developing and middle-income countries, success was measured through outcomes such as: behaviour change and nutritional status, sometimes attitudes and knowledge. These outcomes make up the last stages of a longer process that in the case of nutrition education might or might not result in behaviour change and/or nutritional status change. This view was also supported by Gillespie (1981) 25 years ago[60]: "The focus of most nutrition education research has been on outcome rather than process. Although outcomes (or goals) are an important focus of study, more emphasis needs to be placed on studying the means to that end." [60] The problem in only looking at the last stages of the process is that one might be losing the opportunity to intervene when the most critical changes are occurring and which otherwise might have contributed to a successful impact. But even when there is no successful learning, by looking at the intermediate steps, one can perhaps understand what went wrong in the process. If this is not measured, the failure may be erroneously attributed to the content of the message or other factors.

Hence, one should carry out a process evaluation which would allow us to understand the mechanism of the communication persuasive steps which precede behaviour change and which would help us detect problems in the earlier steps of the process. Overall, this would aid in the design of better interventions.

The field of communication, especially persuasive communication has a lot to offer in explaining this process.

Persuasion has been defined as “any message that is intended to shape, reinforce, or change the responses of another, or others.”[61] It is the communicative act that persuades an audience, implying that it has understood a speech, and believed its message.[62] While nutrition education might only consist of provision of information, as seen in the definition of nutrition education, the aim is not to just give information to the audience of the education program, but to give them such information that will convince them to perform an action. Also, as mentioned by Gillespie (1987), the focus on a nutrition communication, is how people respond to a communication and how nutrition educators can predict and influence people’s responses. There is, therefore, an element of persuasion in nutrition education and hence, nutrition education can be considered as a persuasive communication.[63]

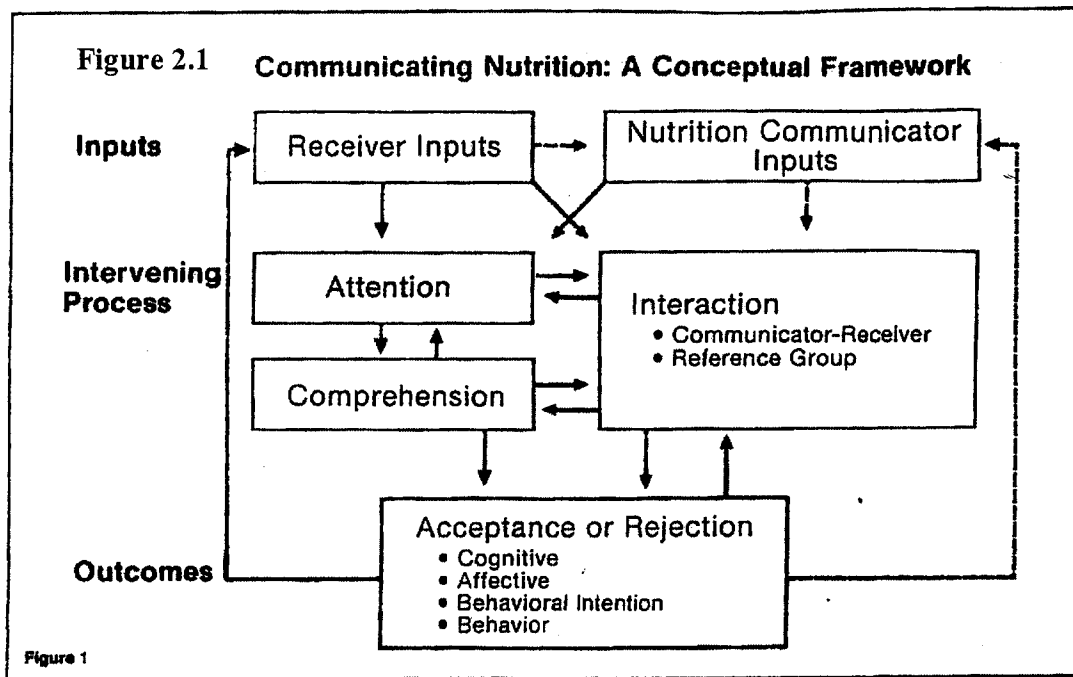
Looking at measuring persuasive communication in a nutrition education program may complete what has been missing in the current evaluation strategies. The best way forward would be to use a model or a theory as a framework for the evaluation. The importance of using a theory is reiterated by Gillespie (1981): “in spite of the advances in evaluation, many nutrition education studies still lack an explicit theoretical base. In order for nutrition education research to advance our understanding significantly, studies need to be designed to include the development and/or application of explicit theoretical frameworks or models which can be tested and improved upon and from which a body of generalizations may emerge to guide further research.”[60]

When searching for the right model/theory it is important to know what is it that we are trying to look at.

Gillespie and Yarbrough have been the main persons linking communication theory and persuasive communication to nutrition education research. Persuasive communication looks closely at two sides, the message side and the receiver side. Based on this,

Gillespie and Yarbrough (1984, 1987)[63, 64] designed a model (see figure 2.1), the nutrition communication model, to evaluate nutrition education programs. The model is made up in one side by inputs and in the other, by responses. In the input side we have the nutrition communicator inputs which control the communication in terms of who sends the message, which channel is used to convey the message and what is the content of the message. At the same time, the receiver itself receives the communication with certain predispositions (receiver inputs). He/she will have certain attitudes and beliefs that might have an effect on the way that he/she responds to the message. Also, the receiver's situational context, for example, her/his socio-economic status will have a role to play. The communicator has little control over the receiver inputs. These two types of inputs combined influence the receiver's response to a communication. Responses occur at two levels: the intervening process and outcomes. In the intervening process, receivers have to pay attention, and comprehend the message and also interact with the message/communicator/fellow peers. This intervening process should then lead to the desired outcome which is the receiver's acceptance of the message on a cognitive level (i.e. believing it to be true), affective level (i.e. attitude towards the message) and behavioural level (i.e. behaviour change).[63, 64] This model, will explain how and why people respond in certain ways to particular communications, why and how people change their knowledge, attitudes, and practices. As mentioned by Gillespie (1987) "receivers are the key element in the process because they can choose to attend or not attend to messages they encounter and they accept or reject the message." [63]

Figure 2.1 The nutrition communication model [63, 64]



It is therefore important that theories/models provide a way of looking at the message and receiver's side, but especially at the receiver's side. Since one of the study aims of this study is to compare two nutrition education interventions, the chosen model/theory has to be able to measure and compare, to some degree, the level of persuasiveness between the two interventions. In other words, this means that the study wants to see what is the effect of different delivery methods (intervention I versus intervention II) on the responses from the receivers. However, the model/theory has to be clear enough and descriptive enough that a series of responses can be a good proxy of the level of persuasiveness of the communication.

Carl Hovland (1959) and his associates at Yale University were pioneers in the study of communication and especially, persuasion from a psychological viewpoint and since then, there have been many theories of persuasion communication that can be applied to health promotion and hence, nutrition education programs. [65] Main ones are for example, the Elaboration-likelihood Model, the Heuristic model, the Theory of Reasoned Action and the Social Cognitive Theory.[66, 67] Many of these are also information-processing theories, which explain how receivers evaluate and process (new) information.[63] Apart from these ones, there is also the Health Belief model and the transtheoretical model, which like the

Theory of Reasoned Action and the Social Cognitive Theory, explain health behaviour and health behaviour change by focusing more on the individual. [53]

While all the above models explore different constructs and some are very good as they contribute substantially to our understanding of individual behaviour and others put individuals into the broader context in which they are living, including many factors that might influence health (including the competence of the community to address the social, economic and environmental determinants of health which are beyond the control of an individual), none seems to be a good candidate for the process evaluation of the TSNi nutrition education program, since none describes in the detail the learning process of those mothers that were exposed to the program. Even though, as will be seen later in this chapter, William McGuire's Communication/Persuasion model has not been based substantial empirical testing, while the above referred models/theories have been used extensively, McGuire's model still presents the most comprehensive description of the relevant variables that make up the outcome of a persuasive communication situation. Apart from that, it is designed as very practical guide and an effective guidance to those wanting to develop communication strategies and then evaluate them. None of the above models/theories look at communication components, for example, at how health messages can be most effectively communicated and acted upon. As Flay *et al* describe: "The most well developed and researched model of the communication process is the classic information-processing model of general attitude and behaviour change first developed at Yale by Hovland in the 1940s and 50s for studying persuasion and has been elaborated upon and extensively advocated by McGuire (1968)."[68]

This research is looking closely at the "how people respond to the communication" side and McGuire's model more than all the other models describes this process with most detail, making it easier to measure persuasiveness in the both group and individual classes and hence, also compare the level of persuasiveness between the two intervention groups.

2.2 McGuire's Communication/persuasion model

This model was developed by William McGuire to design and guide public education campaigns, and is now a commonly used framework for communication planning.[53, 69, 70] This model has also been given other names, such as: "Communication-behaviour change model", "Communication/Persuasion matrix", "Information-processing for behaviour change theory" and "Information-processing theory".

The model is based on communication inputs and outputs, where the inputs include the various components out of which one can construct the communication to change outputs, which consist of the successive information processing behavioural steps that the communication must evoke in the target person for the persuasive impact to occur.[69, 71, 72] These input and output steps are fundamental to communication development, corresponding directly to Lasswell's [1948] well-known description of communication as *who*, says *what*, through which *channel*, to *whom*, and with *what effect* [53, 69, 70, 72-75]

The five communication inputs are:

Source: the person, group or organization from whom a message is perceived to have come. The source can influence the credibility, clarity and relevance of a message.[53, 72]

Message: what is said and how it is said. The content and form of a message can influence audience response. For example the use of fear or humour to communicate the same message may provoke different responses from different target audiences. Practical considerations such as length of message, form of language and tone of voice are also included here.[53] The material used is also part of the message characteristics, for example, using pictures. [72, 76]

Channel: the medium through which a message is delivered. These media include, television, radio, newspapers, direct mail, and more recently, electronic communication.[53] But it also includes the mode of message delivering, if it is a face-to-face communication or mass communication and if it is individually or in a group face-to-face communication. Issues to be considered here include the potential reach of different media, the cost of use, and differences in the complexity of message which can be communicated through different media. [53, 72]

Receiver: the intended target audience. The gender, age, ethnic background, current attitudes and behaviours of relevance, and media use of the target population are all of importance in matching the right message to the right channel from the right source. [53, 72]

Destination: the desired outcome to the communication. This may include change in attitudes or beliefs, or more likely, changes in behaviour. [53, 72]

Variations in the above communication input steps have an impact on the persuasion process by affecting the 11-step sequence of events, representing outputs from a communication, which link initial exposure to a communication to long-term change in behaviour. [53] These are listed below:

- **Exposure:** *was the subject exposed to the message?*
- **Attention:** *Did the subject pay attention to the message?*
- **Interest or liking:** *Did the subject find the message interesting?*
- **Understanding/ Comprehending:** *Did the subject understand the message?*
- **Skills acquisition:** *Did the subject learn how to carry out the relevant task?*
- **Attitude change:** *Did the subject change his/her attitudes?*
- **Memorisation:** *Did the subject retained the message after some time?*
- **Recall:** *Did the subject manage to retrieve the information after some time?*
- **Decision-making:** *Did the subject decide to act based on the information received?*
- **Behaviour change:** *Did the subject change his/her behaviour according to the information received?*
- **Maintenance** – *are subjects still able to maintain the taught behaviours?*

Briefly, the above steps would then fit the process in the following way: the message recipient first has to be exposed to the message. If he/she is indeed exposed, he/she would then have to pay attention to that message. After paying attention to the message, the recipient has to comprehend the message, the overall position it advocates and the arguments provided to support this position. After understanding the persuasive message, the recipient must agree with the message and consequently change his/her attitude. And, if this change is to persist over a period of time, the message recipient must retain, or store in memory, and be able to retrieve his/her changed attitude. Because, the ultimate goal in

a persuasive communication is to lead the recipient to behave in a certain way, the last stage, as a consequence of the described process, would be behaviour change. [70, 72, 77]

McGuire argues that for the recipient to succeed in one step, he/she has to succeed in the step that precedes it. If the sequence is broken at any step, the subsequent steps do not occur. McGuire states that to maximise persuasiveness and lead to a behaviour change, a communication should be designed to maximize effectiveness at each step but he also makes it clear that the evaluation of a campaign must be conducted with regard to the processing step one wishes to achieve (i.e. it could be comprehension of the communication). [70, 77]

McGuire's theory, with its stair-step sequence of responses is complete and descriptive, and indeed, very useful in the creation and evaluation of persuasive communications. To evaluate an existing campaign, one should analyze the input factors effect on the above output steps.[73]

There are some problems with this model/theory. McGuire himself has noted that the model expresses the human mind's response to persuasion as something similar to the internal processing of a computer, but it appears that the cognitive processes within the human mind are not always so linear.[77]

Rimmer and Glassman (1984) argue in their experimental study that " the path is clearly more complicated than a straight line connecting all the variables...There appears to be clear connections among the six outcomes (i.e. exposure, comprehension, retention, belief, attitude and behaviour)".[78]

Another problem, is that it is sometimes difficult to distinguish between the output steps in terms of how they are to be operationally defined, and McGuire himself has suggested that the division points between steps are arbitrary and some "simplifications" of the model can be made for use in empirical studies. Moreover, a program might still be effective without the subject being guided through each and every step of the process. [77]

Also, a change early in the sequence may not lead to a change in later steps because some steps in the sequence may be independent of each other. For example, there is little evidence that comprehending, memorising and retrieving a message will lead to attitude

change and also, attitudes can change even when there is no comprehension, memorisation and retrieval of message (i.e. comprehension, memorisation and retrieval of message are necessary but not sufficient).[79]

Another point that McGuire reiterates is that the model dictates that the probability of any step occurring is proportional to the joint probability that all previous steps occur. This model shows that one should expect campaigns to have attenuated persuasion effects. For example, if the conditional probability of success at each step is 0.80 (since it is unlikely that someone has perfect attention) only 7% of one's target audience would make it through step 12 ($0.80^{12} = 0.069$). Because it would be unreasonable to expect any information campaign to achieve an 80% success rate at every information processing step, most campaigns will have attenuated effects on their audiences. [77, 79]

2.2.1 Systematic Literature Review

Objectives

The main objective of this literature review was to critically look at other studies where McGuire's output step model was also used as a framework to measure and evaluate the impact of health communications. This review will look specifically at:

- Output steps most commonly evaluated
- Methods used and mode of data collection

Methods

Search strategy

Keywords, titles and abstracts of the 5 databases listed below were searched using the various search terms. Google was also searched for grey literature.

Additionally, the reference sections of the relevant articles located through the database/Google searches were hand-checked to identify any additional studies. Also, related articles retrieved from the databases were looked through for any relevant studies.

Authors and known experts in the field were contacted to identify additional published or unpublished studies.

Databases used:

- PubMed
- PsychInfo
- PsychArticles
- EMBASE
- Web of Knowledge (including ISI Proceedings)

Search terms

Due to the area of research, the search terms had to be related to health education, to cognition, to McGuire's model and evaluation.

The appropriate terms for the above words were retrieved from the database's thesaurus.

The following were the terms retrieved from the thesaurus:

- *Health education/ health promotion/Persuasive communication*
- *Mental processes/ cognitive processes/cognition*
- *Attention*
- *Comprehension*
- *Evaluation/ program evaluation*

The following were the various search terms used:

- "health education[MeSH] OR "persuasive communication"[MeSH] OR "health promotion"[MeSH] AND "mental processes"[MeSH] AND "Program evaluation"[MeSH] - [limited to 19+years]
- "McGuire's communication/persuasion matrix"
- "McGuire's communication/persuasion model"
- McGuire's model
- "McGuire's information-processing theory"
- "McGuire's input/output model"
- "health communication" AND "comprehension"
- "persuasive communication" AND "health education" AND "comprehension"
- (model or theory) AND (McGuire's OR McGuire) AND (communication OR input OR output OR persuasion)

The abstracts (and any available full texts) of all studies retrieved with the search terms were reviewed and studies potentially eligible for inclusion.

Those abstracts found to be irrelevant were discarded and those of potential relevancy were further looked at by collecting the full text. Relevancy was defined as meeting the first 3 criteria points mentioned below. Regarding the third point, if it did not mention McGuire's name but mentioned some of the output steps, that were not necessarily knowledge, attitude change and behaviour change, then the full text would be considered relevant.

Inclusion and exclusion criteria

As stated in the objectives to this review, studies could only be included if they used McGuire's output steps as a method of evaluation of a health education program. Publications identified via the search processes were required to meet the following criteria:

- (i) Studies had to be intervention studies but not necessarily RCT.
- (ii) Subjects had to be adults (18+ years). The main reason for this is that results are to be extrapolated to the research study subjects which are also adults.
- (iii) Studies had to use McGuire's output steps to evaluate the health campaign and had to clearly say it.
- (iv) Studies had to include a results section

No restrictions were placed on date, geographical origin or language of publication.

Papers were not excluded on the grounds of methodological quality. Methodological limitations were evaluated, documented, and are presented as part of the review.

Analysis of data

The methods used to measure the different output steps will be mentioned instead in the methods chapter under the relevant section. The methodological limitations of the included studies were assessed in order to give an indication of the validity of the study results. This involved evaluating the following factors:

- Output steps reported to have been used and if they are equivalent to those mentioned in McGuire's model
- If validation and reliability of methods was measured

- If randomisation during selection of participants was reported or if there is a selection bias

Greenhalgh in a publication in the BMJ[80] suggests looking as well at the following points:

- Was the study original?
- Whom is the study about?
- Was the design of the study sensible?
- Was systematic bias avoided or minimised?
- Was the study large enough, and continued for long enough, to make the results credible?

Results

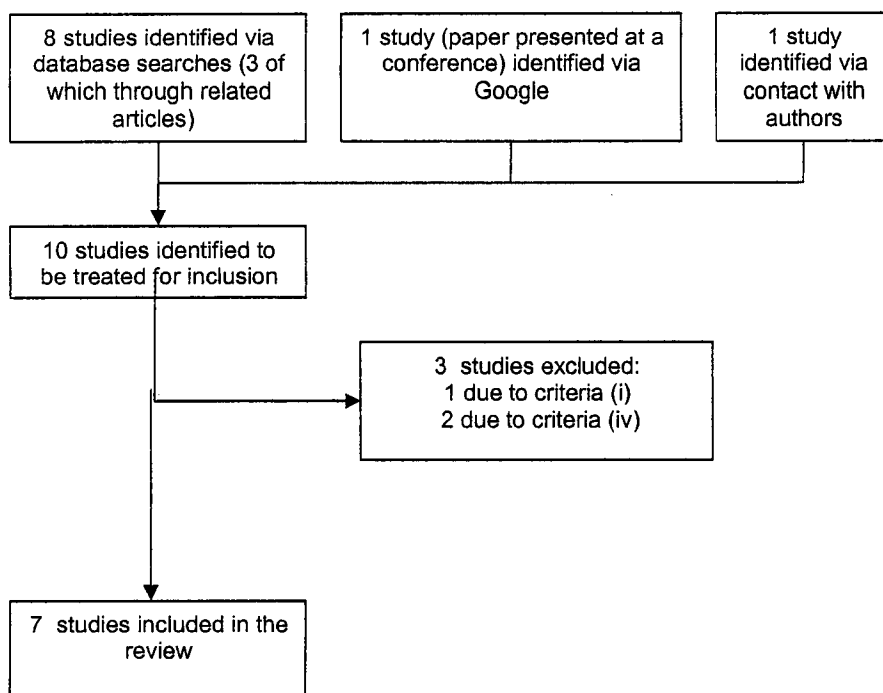
Selection of studies

The search terms resulted in 154 hits in Pubmed, 256 in Psycinfo, 168 in EMBASE, 3 in the Web of Knowledge, and 33 in Psycharticles. These hits were then refined following the search criteria and the objectives of this review.

As can be seen in figure 2.2, a total of 10 studies were identified as a result of the search process outlined above. Eight of these were located via databases (three being related articles), another through a search in Google and another through personal communication with the author.

One of the 10 publications, a review by Houts[81], was excluded since it used McGuire's output-step model as the framework for the review but did not necessarily review studies that have used McGuire's output-step model. The study retrieved through personal communication with the author John Elder[82] will also not be included in this session as it is only a draft for a proposal and does not include results. The third study not to be included is a French study by Hausser[83], again a proposal for a model of evaluation that does not include results. Nonetheless, the methods used to measure the output steps in each of these studies which were relevant to this research study will be critically looked at in the methods chapter.

Figure 2.2. Selection of studies



The included studies

All of the 7 included studies were set in high-income countries (USA and Australia). One study was a conference proceedings and the remaining 6 were published in journals. The earliest publication date was 1984 and the latest was 2006.

All of the studies used the model to evaluate the following types of health education: “printed health education materials”, “tailored women’s health magazine”, “persuasive health messages”, “mass media campaign”, “evaluation report”, “nutrition education program” and “AIDS prevention ads”. The health behaviours aiming to be achieved by the studies were: AIDS prevention, asthma management, exercise promotion, cancer screening and prevention, weight loss, nutrition related behaviours (i.e. calcium, baby bottle tooth decay, vitamin A, food safety and immunization) and agriculture extension.

Table 2.1. Description of studies that have used McGuire's communication/persuasion model to evaluate a health education program

Study ref	Justification for the use of McGuire's model	Output steps measured	Validation and reliability?	Study design, sample size and participant details	Mode of data collection	Results
Rimer (1984)[78]	"To test an empirical definition of comprehension and to investigate the path leading from comprehension to subsequent retention, attitudes, beliefs, and behaviour change"	1. Comprehension 2. Retention 3. Attitudes 4. Beliefs (not part of McGuire's output steps) 5. Behaviour change	Validity NR; Reliability reported for comprehension scales	-Post-test only -Control group included - Addresses in Maryland were selected randomly -Participants 18 yrs and older from both sexes - Sample size = 123 respondents - Sample size in each group not reported	Two 10 min interviews (second interview was telephone based)	Intervention subjects performed better in the comprehension score and retention score; Exercise increased significantly in both intervention and control subjects; Results for attitudes and beliefs are not presented Comprehension (operational and psychological meaning but not lexical meaning) was significantly correlated with attitudes, retention and self-reported behaviour
Navaie (1994)[84]	"McGuire's information processing model was used to compare content of materials with comprehension, retention, and yielding of the participants"	1. Comprehension 2. Retention 3. Yielding (or described as attitude change by	NR	-Cross-sectional design -Participants	Focus groups	Simple messages without technical terms were better comprehended (i.e. less difficulty in

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Comino (1997)[86]	"To present the results of evaluation of these campaigns upon asthma awareness, knowledge, and management in the Australian population"	6. Yielding (or attitude change by McGuire) 7. Action (or described as behaviour change by McGuire)		UIUC and 42 senior graduate students from the University of Americas Randomisation is not reported Sample size calculations not reported	Telephone surveys self-report	Factor analyses was used to find which items related well with the output steps. A matrix was created for all those items that were related to specific output steps.
		1. Awareness (or described as exposure by McGuire) 2. Recall 3. Knowledge (or described as comprehension by McGuire) 4. Management (or described as behaviour change by McGuire)	NR	--Cross sectional design -participants mixed sex, 18 yrs and over, selected from the general pop -randomisation of participants reported -- sample size calculations not reported		Overall an increase trend in awareness (χ^2 trend =263.4, $p<0.001$), recall (χ^2 trend =301.9, $p<0.001$) and knowledge ($p<0.01$)

Bull (2001)[87]	"The purpose of this study was to examine the relationship between the communication inputs and the cognitive and behavioural output using an adapted version of McGuire's communication persuasion model."	<ol style="list-style-type: none">1. Exposure2. Attention3. Liking4. Understanding5. Skills acquisition6. Yielding (or described as attitude change by McGuire)7. <i>Memory storage/agreement</i>8. Retrieval9. Decision-making10. Behaviour change.	NR	Follow-up study with baseline and post study measurements Participants: BMI>27kg/m2, 18yrs or more, Mixed race and mixed sexes, with an expressed interest in losing weight Randomisation of participants not reported Sample size: 198 No sample size calculation reported	Questionnaire: Baseline, post intervention and 1 month follow-up (telephone interview). Self-report	Results are not results for output steps per se, but results for their association with input steps. All output steps were positively associated with some of the inputs variables.
Kreuter (2004)[88]	Assesses women's responses, based on McGuire's communication/persuasion model, to a tailored cancer communication magazine	<ol style="list-style-type: none">1. Attention2. Liking3. Yielding (or described as attitude change by	NR	--Cross sectional --Included a control group --1227 African	Interviews (baseline interview, 1- and 6-month follow up telephone	There was no difference in attention scores (Anova, p=0.35), liking (Anova, p=0.12),

			McGuire) 4. Memory storage 5. Relevance (not part of McGuire's output steps)		American women ages 18-65 randomisation of participants not reported	interviews). Self-report	attitudes (Anova, p=0.99) between the 3 study groups. Memory storage was the only output step that showed a significance difference between the 3 study groups (Anova, p=0.04). Did not compare study groups with control group.
Zhao (2005)[89]	This study examined the effectiveness of ads using fear appeal on attention, liking, interest and intent to behave.	1. Attention 2. Liking 3. Interest 4. Intention to behave (or described as decision making by McGuire)	Reliability reported for interest scales Attention and liking items were borrowed from Bull's study;	--Repeated measures (within subjects) design Randomisation to each of the experimental conditions Sample size calculation not reported Subjects undergraduate students Sample size: Male high fear	Questionnaire Self-report	There was no significant difference in liking, interest and decision making between the within-subject repeated measures. However, attention was higher for high fear PSAs (t-test, p<0.001)	

				<p>appeal =48</p> <p>Male low fear appeal = 42</p> <p>Female high fear appeal = 53</p> <p>Female low fear appeal = 44</p>			
--	--	--	--	---	--	--	--

NR – not reported

Type of study and study design

All of the studies, except Russon's study[85], were health education interventions. Russon's study does not specify what kind of topic it was, but it was assumed in this study to be agriculture since subjects were graduate students from the department of agriculture studies.

All of the studies used McGuire's output-step model to evaluate receiver's reaction to the program message or report. Only 1 study[78] reported using a control group, and 4 studies[78, 86-88] reported randomly selecting their subjects. The studies that did not use a control group either used pre- and post-tests on the same subjects or just used post intervention tests.

All of the studies measured the output steps at the same time, right after subjects were exposed to the messages. The exception to this is Rimer's[78] study that measured retention (memorisation) 4 weeks after the intervention. The advantage of measuring the output steps at the same time is that measurements for each output step can be made for the same subjects. Moreover, the researcher only needs to administer the tests once.

The use of McGuire's model could be divided into those studies that only wanted to look at the impact of the intervention on the output steps (i.e. wanted to measure the output steps per se), therefore focusing more on the results for the output steps, for example Comino's and Rimer's studies, with the latter study even exploring one of the output steps, comprehension. And those studies that wanted to explore the influence of independent variables (i.e. the input steps), and their characteristics, upon the communication/persuasion process (i.e. on the output steps). For example, Kreuter's study wanted to measure responses (measured through results for output steps) to messages that used either behavioural and socio cultural constructs (input steps), and Zao's study wanted to measure responses to different degrees of fear in messages, and Navaie's study measured effect of message characteristics on comprehension, retention and attitude change. Furthermore, Bull's and Russon's studies examined the relationship between a series of inputs and outputs.

Output steps evaluated

As can be seen from the table, some of the output steps measured were termed differently from McGuire's model, and they have also added other constructs that do not come up in the model. Comino's[86] study reported to have used constructs from McGuire's model, even though awareness and knowledge are not terms used for any of McGuire's output steps. Looking carefully at Comino's constructs it seems that awareness would be the equivalent to exposure and knowledge would be their equivalent of comprehension. Rimer's[78] study reported to have used an adapted form of McGuire's model and hence justifying for the use of the construct beliefs which is not part of McGuire's model. Kreuter[88], also measured relevance which is not part of McGuire's model. Zhao's[89] study measured intention to behave and it was not clear to each output step was intention to behave equivalent to. In this review it is considered as decision making, however, it could also be attitude. Navaie's[84] study measured comprehension including retention of messages and their short-term recall. This seems more a measurement of retention (memorisation) and recall than actually comprehension, and indeed the authors themselves wrote: "perhaps future studies may be better able to confirm these findings by directly determining whether messages that are retained by the audience are truly comprehended and not simply repeated." All of this brings about confusion on which terms the authors were trying to measure.

Overall, studies measured between 3 to 10 output steps, with Bull's[87] study measuring the most. Only two studies measured exposure, but one seemed to call it awareness and another considered it as 100%, therefore not really measuring it. Four studies measured attention, 4 measured liking, 1 measured interest, 5 measured comprehension (even though one called it knowledge), 5 measured either memorisation or recall, 5 measured attitude (change), 2 measured skills acquisition, 2 decision making and 4 behaviour change (even though one called it management). So as can be seen, throughout the studies there is no consistency on the number of output steps measured.

Methods and mode of data collection

Different methods were used to measure the same outcome steps, even though this is not described in the above table. This difference probably comes from different operational definitions used among the studies. For example in Russon's study the definition of comprehension was "the absence of qualities within an evaluation report that would cause the receiver to misunderstand the report" while in Navaie's study it was "the selective retention of messages, their short-term recall and their articulation into speech" and in

Rimer's study "the shared understanding of the psychological meaning, operational meaning, and lexical meaning of a health message". Regarding comprehension measurements for example, one study[86] assessed level of agreement/disagreement with a statement, another[87] asked subjects how easy it was to understand the message, and another[84] would compare recall of messages with the actual messages in the health material. Regarding attention measurements, studies either asked participants to rate the degree to which the message caught their attention, or asked respondents to rate the amount of attention that they paid, or asked respondents if they had read the message or asked them question related to the influence that the message characteristic might have had on the degree of attention. The same happened for all the other output steps, showing that there is not just one method to measure them. Moreover, it is possible that studies are not using the most appropriate methods to measure them. Rimer[78], in her study measures two types of comprehension, and explains and shows in her study that even though comprehension lexical meaning, which is made up of recall and recognition questions which measure what people say, is the type most frequently assessed in health education programs, it does not correlate with behaviour change. While comprehension psychological-operational meaning, which measure what people can do, correlates well with behaviour change. Her study was the only one that measured the second type of comprehension.

The only aspect in common between the studies, regarding the methods used to measure the output steps, is the fact that the majority of the studies used self-reports derived after exposure to the messages. An example of this: "Did you find it easy to understand the material?"[87]. A major limitation of these explicit measures is that they require introspection from the participants and, participants may respond with socially desirable statements.[90]

Quality of the studies

The assessment of the quality of the studies used in this systematic review was based on a set of questions posed by Greenhalgh [80].

Quality was also further assessed by checking if studies reported the validity and reliability of methods used and randomisation of participants.

Greenhalgh's first question is: *Does this new research add to the literature in any way?* To look at this point she suggests a series of other points:

- Is this study bigger, continued for longer, or otherwise more substantial than the previous one(s)?
- Is the methodology of this study any more rigorous (in particular, does it address any specific methodological criticisms of previous studies)?
- Will the numerical results of this study add significantly to a meta-analysis of previous studies?
- Is the population that was studied different in any way (has the study looked at different ages, sex, or ethnic groups than previous studies)?
- Is the clinical issue addressed of sufficient importance, and is there sufficient doubt in the minds of the public or key decision makers, to make new evidence "politically" desirable even when it is not strictly scientifically necessary?

Population assessed in all different studies varied since the communication messages were targeting different populations. For example, Kreuter's study targeted African-American women in waiting rooms. Because of this, it might not be possible to generalise findings to other populations. None of the studies address specific methodological criticisms of previous studies. Comino's study and Kreuter's study were the biggest study, with 1800 and 1227 participants respectively. They were also the longer studies running for 6 months. All the other studies were small and measurements were taken right after the exposure to the messages (some studies followed up participants for a month after exposure to the communication), hence, duration was short. Nonetheless, the outcomes measured did not need a longer study duration. In a way, all studies are unique as some measured different output steps than others, and only 1 measured them all. Also, some measured responses to specific message characteristics and some measured to all input steps. As well, they all vary in the methods used. So, they all in their way, contribute with new information to this subject. All studies include a discussion of weaknesses in the studies and how these might have limited their findings, they also discuss implications of their findings.

Greenhalgh's second question is: *"Whom is the study about?"* This point looks at: how were subjects recruited, who was included in the study, who was excluded from the study and if subjects were studied in a real life circumstance.

In all studies, subjects were studied in a real life circumstance. Only two studies randomly selected their participants (Comino[86] and Rimer[78]), hence, ruling out the possibility of selection bias. Of those that did not, two (Russon[85], Zao[89], Kreuter[88], and Navaie[84]) were evaluating a campaign/program/material in a specific population and therefore had to select participants from that population (i.e. graduate students from specific faculty, or a specific clinic attendees). Bull's study[87] was the only one that could have incurred some selection bias since study participants responded to a newspaper advertisement that invited participants to join the study. However, a selection criteria list was used to then screen participants. This screening ruled out the possibility of selection bias as all potential risk factors were accounted for.

All the studies, except Comino's, proposed to measure something very specific which meant as well measuring it in a specific population. This also meant that strict selection criteria list was not used. Only Kreuter and Bull's study used selection criteria to select participants. Regarding Kreuter's study one cannot know if results can be generalised to a population that did not have a telephone and could not read at a fifth grade level since these were some of the excluding criteria.

Greenhalgh's third question is: *"Was the design of the study sensible?"* This point looks at: What specific intervention was being considered, and what was it being compared with? What outcome was measured, and how?

The above latter question has already been dealt with previously under output steps evaluated and methods of data collection and hence will not be repeated here.

All studies used different designs and all seemed sensible for the type of study. None pretended to be a RCT or to use similar rigorous designs because there was no need.

Greenhalgh's fourth question is: *"Was systematic bias avoided?"* This point is only applicable for studies that are comparing groups. Only two of the studies compared groups of some sort. Rimer's study compared an experimental group to a control group and Zao's study included 4 experimental conditions. The groups should be as similar as possible except for the particular difference being examined. In the case of Rimer's study, participants were randomly selected from a sampling frame and then still further randomised to either experimental or control group. This should have been enough to ascertain that the two groups should be relatively similar. In the case of Zao's study

participants were also randomly allocated to the 4 experimental conditions. However, it is not clear in the study if participants were also randomly selected. If that is not the case then we do not know if results can be generalised to a wide population of undergraduates.

Greenhalgh's fifth question is: "*Was assessment blind?*"

This is again an important factor for studies that are comparing groups, especially when one of the groups is a control group. If the person assessing is not blind to the group status then there might be a performance bias. Neither of the above studies (Zao's and Rimer's) that compared groups was blinded to the groups status which could have affected results, meaning that, neither the participants nor the assessors were blind to allocation.

Greenhalgh's final question is: "*Were preliminary statistical questions dealt with?*" This includes points such as: the size of the sample; the duration of follow up; and the completeness of follow up.

None of the studies reported any sample size calculations and the importance of this is that studies could have been underpowered, leading to a type II or β error - the erroneous conclusion that an intervention has no effect. A study should be big enough to have a high chance of detecting, as statistically significant, a worthwhile effect if it exists, and thus to be reasonably sure that no benefit exists if it is not found in the study. A sample size calculation would have ensured that the study was big enough.

Duration of follow up was not an important factor for these studies since all of them wanted to measure the impact of a communication and depending on the outcome it could be straight after the communication or sometime afterwards.

It is not clear from the studies if the data from those subjects that dropped out from the study was also analysed.

Only one study reported carrying out a validation of the methods used, another study reported that the items were borrowed from a study, and 3 studies reported carrying out some type of reliability test.

From this review it can be concluded that only a few studies have used McGuire's output step model to evaluate health campaigns and all were carried out in developed countries. No studies have made use of the model to guide an evaluation of a health program in a

developing country. One can also conclude that there is not a specific number of output steps that are usually measured, and there is not only one way of measuring the output steps which can be seen by the lack of consistency in the methods used through out the studies for the same output steps. Studies failed to report on several issues that are used to measure their quality. For example, they do not consider the validity and reliability of their methods and tools, and also, they do not report any sample size calculation for the effect expected with the communication. Some studies were also carried out in a specific population and hence is not representative of the general population. The review also shows that studies tend to measure the output steps using self-reports, which can lead to subjective responses whose validity can be questioned.

All of these conclusions are taken into consideration in chapter 4 which describes the output steps measured and the methods and tools used to measure them.

The current study decided to use an adapted and simplified 6-output step version of McGuire's model leaving behind the constructs: interest/liking, memorisation, decision-making, behaviour change and maintenance. Recall was measured instead of memorisation since for the mothers to be able to recall would first have to memorise and as will be seen in chapter 4, one measure of memory is ability to recall or retrieve information. Interest/liking and decision -making were also not measured because self-reports, which is the method usually used to measure them would retrieve very subjective but especially biased answers.. The researcher doubted that any of the mothers would have said that they did not like or did not find the class interesting. Moreover, as was seen from the TSNi individual counselling visits, all mothers agreed that they would try to adhere to the 12 recommendations when in fact they were not. Maintenance was not measured simply because the study did not run for long enough to measure it.

2.3 Literature review of nutrition education studies comparing both approaches and interventions

A major aim of the current research study is to use McGuire's output steps and compare them between two intervention groups.

Even though the current study is not strictly comparing group with individual class approach, these two approaches of communication imply that mothers are receiving the communication in different media which might possibly lead to different effects. For this reason a literature review was carried out on nutrition education studies that have either compared both approaches, or have used both interventions.

Mensing and Norris (2003)[91], reviewed studies that compared group and individual education. Of the 4 studies that compared both approaches, one showed a positive difference for some outcomes (i.e. diet and physical activity) for the group class approach and no difference for other outcomes (i.e. learning self-care skills), another showed a positive difference for outcomes (i.e. level of activity and asthma status) for the group class approach, but the other two showed no significant differences in outcomes (improvement in knowledge, change in blood pressure and weight, reductions in HbA1 and BMI) for both approaches. [91] They also reviewed a meta-analysis of the effectiveness of diabetes education and found that there was no significant difference on the predictor of glycemic control between the two approaches. [91] However, they do not mention if the delivery was collaborative or didactic and what was the group size. Moreover, none of the studies were carried out in a developing country.

A search was carried out in Google and PubMed using the following search terms: "approaches to nutrition education"; "group approach AND individual approach AND nutrition education"; "comparing group class AND individual counselling in nutrition education"; "group versus individual nutrition education"; "group versus individual sessions"; "individual counselling/one-to-one approach versus group class/ group counselling".

13 studies were retrieved and are described in table 2.2.

Of the studies retrieved, one (Rickheim *et al* [92]) was already described as part of the above Mensing and Norris review. It is very important when doing reviews, not to double count studies.

The key points that emerged from this review is that, overall studies showed that there is not much difference in outcomes when group class approach is compared to the one-to-one approach and that both approaches are as effective, even though more effective in different practices measured. Only the study by Hakala *et al* [97] showed that individual counselling lead to more sustained weight loss.

We do not know what would be the outcome if instead the one to one approach was an additional approach to a group class approach. Siero and Rodriguez-Garcia studies [94, 95] were the only ones that made this comparison, however, Rodriguez-Garcia study only presents preliminary results which do not compare the two intervention groups. Siero's study reports no difference between the intervention groups, however, the group allocated to receive the information individually, received it through the mail and not face-to-face like the TSNI study.

The studies reviewed only compared the last stages of the communication process, between the approaches and intervention groups. This study adds to the current gap in the literature by comparing the earlier stages between the intervention groups.

Moreover, apart from 2 studies[95, 96], all of the above studies were carried out in a high income country where processes within the approaches may be quite different. For example, education is much less didactic, group class approaches are mainly counselling led by highly qualified staff, and group sizes tend to be much smaller. Thus, this research study also adds to the gap in the literature in that earlier steps of the communication process were measured in the contrasting environment of a developing country setting.

The quality of the above mentioned studies was measured by looking at the following questions: did the studies explain clearly the objectives and did they attempt to measure them? Were the methods used clearly explained, including participant selection? Were they adequate for the study? Was data analysis described and was it adequate? Did they refer to any weaknesses/limitations and their implications for the study?

All studies included a clear description of the objectives and attempted to measure them. Methods were always clearly described, including participant selection with a description of exclusion/inclusion criteria. The majority of studies randomly assigned subjects to comparison groups. However, none of the studies reported carrying out sample size

calculations and indeed in 1 study[97], sample size was very small and in another[98] a big discrepancy in sample size between the group and individual approach was observed. Also, for this same latter study response rates for the individual approach were higher than the group class approach, which might have biased the results. Only one study mentioned reasons for drop outs.[99] Data analysis was explained in detail in all studies and seemed adequate. All studies mentioned weakness and limitations of the study and their implications for the results, for example the study with the small sample size discussed this and that it was not representative of the obese population. All studies initially mention how their results can add to the gaps in the literature since not many studies have compared the two delivery settings.

Table 2.2 Description of studies comparing a group versus individual nutrition education approaches and the two intervention groups

Author	Aim	Design	Results
Hakala et al [93], 1993	To compare group and individual weight reduction programmes in the treatment of severe obesity	30 subjects randomly assigned to group counselling (GC) received an intensive two week education. 30 subjects randomly assigned to individual counselling (IC) were seen once a month. Counselling was carried out by physicians and nutritionists. Weight reduction program was based on a dietary counselling and physical exercise. Group size was 10 subjects.	Adherence: 69% for GC and 86% for IC in 1 st yr Weight loss: mean weight loss in women during the first 3 months was greater in the GC than in the IC ($p < 0.001$) Mean weight loss in men was similar for both groups in the beginning but at 8 months it was greater for the IC ($p < 0.01$) Later follow-ups showed no significant differences between the groups.
Cornier [97], 1972	To determine if a group approach would lead to greater weight loss in comparison to an individual approach	Subjects randomly assigned to either group or individual treatment. Weekly meetings were held for 12 weeks. In both treatments the educational content was identical and lasted approximately 30 min. A dietician led both groups. Group size not reported	There was significant weight loss in the study subjects for both approaches but here was no significant difference in the rate of weight loss between the group and the individual treatment.
Jeffery et al [100], 1983	To compare the effectiveness of a set of educational materials delivered via group treatment and individual counselling.	Subjects randomly assigned to either group treatment format or individual counselling format. Session on weight reduction and sodium restriction. Sessions were staffed by either a nutritionist, health educator, physician, etc. Both approaches received the same sessions. Group size was 7-13.	Significant reduction in blood pressure, weight, urine sodium and dietary measures were seen for both treatment modes from baseline to endpoint evaluation, but no differences were seen between the two treatment modes.
De Mello et al [101], 2004	To compare two strategies for childhood obesity management: ambulatory assistance (individual) and educational program (in group).	Children and adolescents were randomly allocated to two groups: individual outpatient treatment and childhood obesity education program in groups. Lessons given by research assistants Group size not reported	Reduction of body mass index, obesity index and caloric intake was similar in both groups. As for food habits, individual treatment increased the intake of fruits ($p = 0.033$) and vegetables ($p = 0.002$) and reduced the amount of French fries and crisps ($p = 0.041$), while group program reduced the intake of soft drinks ($p = 0.022$), sandwiches, pizza and fast food ($p = 0.006$).
Moe et al [102], 2002	To evaluate the efficacy of two intervention strategies for improving nutrition and physical activity practices in fire fighters.	Randomly assigned fire stations to (a) team-based curriculum, (b) individual counsellor meetings, and (c) usual-care control group. Each except control group was based on a different theoretical model for promoting healthy nutrition practices and regular physical activity. Contact time and educational content were comparable between the two interventions.	Both interventions resulted in significant reductions in LDL cholesterol. The team approach significantly increased personal exercise practices compared to the one-to-one strategy. The one-on-one strategy significantly increased dietary self-monitoring and decreased

Rickheim <i>et al</i> [92], 2002	To compare the effectiveness of delivering diabetes education in either group or individual setting	Counselling given by a health counsellor. Team curriculum was peer-taught 87 subjects assigned to group education (GE) and 83 subjects assigned to individual education (IE). Both groups received 5-7 hrs of education given at different times. Education was carried out by a diabetes nurse specialist and a diabetes nutritionist specialist. Group size was 4-8 subjects	behaviours related to higher dietary fat intake. Adherence: 50% in GE and 59% in IE completed the program. At 6 months, knowledge, weight, BMI, attitude significantly improved for both groups but was not significantly different between both groups. Difference in HbA1c was marginally greater for GE than IE ($p<0.05$). But could be due to medication. Both groups significantly improved level of knowledge, reduced HbA1c, HDL-C, BMI and systolic pressure, but there was no significant difference between the groups.
Llorca <i>et al</i> [103], 2003	To compare the effectiveness of group versus individual health education in type 2 diabetic patients.	Subjects were randomly assigned to group education (GE) or individual education (IE). Both groups received 3 sessions lasting 40 min and similar content. Sessions were given by physicians and nurses Group size was 5 subjects.	Attitudes, subjective norms, perceived behaviour control, and intentions towards nutrition adherence, self-reported nutrition adherence and glycemic control significantly improved in both groups, over the three-month study period. Yet, those receiving individual counselling with group education showed greater improvement in all measures with the exception of glycemic control, where no significant difference was found between the two groups at three months.
Gucciardi <i>et al</i> [104], 2007	To examine the impact of two culturally competent diabetes education methods, individual counselling and individual counselling in conjunction with group education, on nutrition adherence and glycemic control in Portuguese Canadian adults with type 2 diabetes over a three-month period	Randomly assigned to group class and individual counselling intervention or individual counselling only intervention Team consisted of dietitians, endocrinologist, nurses and others. Group class size was 5-8 subjects	Implementation cost was twice as high for the individual intervention. Usage, knowledge and safety of composition of ORT did not differ significantly in both approaches when measured a year after the intervention.
Chowdhury <i>et al</i> [96], 1988	To compare the outcomes (with respect to usage and safety) of two approaches in teaching ORT: individually and in groups.	Female health worker arranged with mothers to come together and taught them some important messages about diarrhoea and ORT preparation and use and then asked the to prepare the ORT under her supervision. The same was done individually to other mothers. Group size was 5 subjects. Setting: Bangladesh	Authors only reported comparison between intervention groups and the control group but did not compare specifically group 1 with group 2
D'Erasmio-Melkus <i>et al</i> 1992[99]	To evaluate the impact of diabetes education between three levels of education intensity	Group 1: 11 group sessions + individual session (for content clarification); Group 2: 11 group sessions + individual session (for content clarification) + 2 individual follow up sessions; Control group – customary medical care Group size not reported	

Adams et al[98], 1986	To compare factors predicting weight loss between a group and individual approach	Group (n=125): 12 weight loss classes lead by dieticians trained in counselling. Class size varied from 8-12 subjects. Individual (n=28): 15 weight loss sessions lead by the same dieticians	Weight change during the program did not differ between group and individual approach. Weight change after the program did not differ between group and individual approach, however, it differed when it results were broken down by sex.
Rodriguez-Garcia et al[95] 1990	To compare the impact of different delivery models for promoting breast-feeding practices	Group 1: Individual counselling with health promoters Group 2: Group teaching with supervisor Group 3: a combination of group teaching and individual counselling Control group: no teachings; Size of group class not reported.	Preliminary results only and do not include results of the comparison between group 2 and 3.
Siero et al[94], 2000	To compare the effect of two interventions focused on the promotion of Mediterranean nutrition behaviour	Group 1: group session where a Mediterranean diet was promoted; Group 2: group session where a Mediterranean diet was promoted plus individual stage-matched education information (mailed to the subjects). Control group – received only a printed leaflet Group size was 10 participants	After 16 weeks both interventions resulted in significant changes in comparison with the control group. Outcomes regarding fish consumption (including attitudes and nutritional intake) improved for both interventions. Also, outcomes regarding fruit and vegetable consumption were also greater for both interventions. But, there was no difference between the intervention in the above outcomes.

2.4 Aims and Research questions

The aims of this research study are the following:

- To define methods/tools to enable the measurement of McGuire's persuasive communication's output steps in a resource poor developing country and to measure their validity and reliability whenever possible
- To determine whether the methods/tools developed can detect variability in response within a group of rural women receiving nutrition education
- To use the methods/tools developed to compare two nutrition education delivery methods in rural Mozambique

The research questions for this research study are the following:

- Were the tools developed able to detect variability within groups for each of the output steps?
- Can McGuire's Communication/Persuasion model detect differences in the communication process between the two Nutrition education interventions?
- Does exposure, attention, understanding, skills acquisition, recall, and attitude scores differ significantly between the intervention groups?
- Does the degree of persuasiveness differ significantly between the intervention groups?

The hypothesis for this research study are the following:

- The various output steps in the persuasive communication process: exposure, attention, understanding, skills acquisition, memorisation/recall, and positive attitude will have higher scores when measured for intervention II mothers (group classes plus individual classes) than for intervention I mothers (group classes only) because of the added benefit of the individual class approach.

CHAPTER 3

THE TSNI STUDY

This research study was an integral part of the Towards Sustainable Nutrition Improvement (TSNI) study. This chapter therefore provides information on the TSNI study objectives, study setting and design, and intervention activities, including a detailed description of the two nutrition education approaches, and their respective processes, as well as putting the TSNI study in the context of the MRC complex interventions description. This chapter also provides information on the measurements collected and an account of published results for the TSNI study.

3.1 TSNI project objectives, study setting and design

The TSNI study was a quasi-experimental cluster randomized, controlled trial research project. It integrated nutritional (i.e. nutrition education), agronomic, and socio-economic components to achieve the project objective: to determine whether a food-based intervention strategy can lead to sustainable, year-round intake of vitamin-A rich foods, reduced fluctuations in seasonal household calorie supply, and an overall improvement of diet diversity, nutritional status and diet quality in a cost-effective manner, particularly among children under five years of age. Emphasis was placed on addressing two major nutritional problems: vitamin A deficiency and inadequate caloric intake. To achieve this, the project introduced pro-vitamin A (beta-carotene) rich varieties of sweet potato (identified by their orange-flesh), utilizing these new cultivars as a low-cost, effective entry point for improving the kinds of weaning foods given to young children and increasing the frequency of intake of essential micro-nutrients and calories. [16] [105]

The following is a description of the sampling procedure for the TSNI study and the rationale for the adopted sampling procedure. The main reason for adopting cluster randomisation was to reduce the implementation cost, especially because the TSNI was only a pilot study, and to avoid contamination between intervention and control groups. [16] [105]

The TSNI study used a multi stage sampling procedure whereby three districts (Mopeia, Namacurra and Nicoadala) in Zambézia Province were purposefully selected (see figure 3.1). Mopeia and Namacurra were selected as the intervention districts mainly due to logistic reasons as the implementing partner (World Vision Mozambique) was operating there, but also, because there was a common dominant language, and it would be feasible for extension staff to travel throughout the districts, also because high levels of malnutrition and vulnerability to drought (and thus potential to benefit) existed there. The above rationale also applied to the choice of control district (Nicoadala). The control district was also purposefully kept separate from the two intervention districts in order to avoid any possible contamination. The next stage, was the choice of localities. Again, two localities were purposefully selected in each of the intervention districts. The reason for their selection was mainly due to logistic reasons as the implementing partner (World Vision Mozambique) was operating there but also, because they had similar agro-ecological conditions. Another 4 localities were purposefully selected in the control district. These control localities were chosen to match the agro-ecological conditions of the intervention localities. However, a prerequisite for their selection was that they could not have been exposed to the OFSP. The third stage was the selection of villages. Villages within intervention and control localities were randomly selected. All the villages within a given agro-ecology were 1st purposely stratified into those within 10 km of a road (where access to services would be easier) and those >10 km of a road. This was so to assure that enough isolated communities were included. Villages that were in major flood zones (next to the Zambezi river) and all villages that did not have year round access by vehicle (i.e. villages only access by canoe during the rains) were purposefully excluded due to logistic reasons/cost. [16] [105]

A critical component of the TSNI study was to determine the most cost-effective strategy for improving dietary quality through dietary diversification and improved feeding practices. Two strategies of nutrition education delivery were compared: Intervention I and Intervention II. For that reason selected villages were then randomly selected and allocated into two sub-group: Intervention I vs intervention II just up to total number of estimated households needed. Intervention villages were selected prior to control villages so that the latter could be matched as closely as possible in terms of agro-ecological conditions. Villages in intervention districts were randomly selected with probability of selection proportional to size. Villages in the control district were also selected with probability of selection proportional to size but were eliminated if OFSP had been distributed in the area or if there were plans for distribution. [16] [105]

In intervention I, community nutritionists held monthly women group nutrition/health sessions; whereas in intervention II, in addition to the group sessions, home visits were made every other month to participating female caregivers in the study. In the final sample, 49% of the principal female caregivers were in the intervention I intervention (246 households), compared to 51% in the intervention II intervention (252 households). Both intervention groups received the same agriculture extension activities. The control group received neither nutrition education activities nor agriculture extension activities. [16] [105]

Overall 39 villages were selected in the intervention areas and 10 in the control areas. Complete household listings were conducted and all households with an age-eligible child (4–38 mo) were invited to participate. In the intervention villages, participation in the study entailed membership in a farmers' group; control villages had no such requirement. Of the 827 households initially enrolled in the study 90% (741 households) completed the study: 498 households in the intervention districts and 243 in the control district. The most common reason for drop-out was due to the death of the reference child (37 cases), closely followed by the household or the child leaving the study area (32 cases). In only 13 cases did families refuse to continue in the study.[105]

Limitation in the TSNi study design affecting the current study:

The TSNi study design is such that output steps can be compared between the two interventions; group classes and group classes together with individual classes. It is not, however, possible to make direct comparisons between the two educational approaches of group or individual classes since everybody who received individual classes also received group classes. However, since the setting in which measurements took place (group class or individual class) is likely to have a substantial impact on the output step, this must be taken into consideration and it may be possible to make some inferences about the different approaches through comparing the interventions.

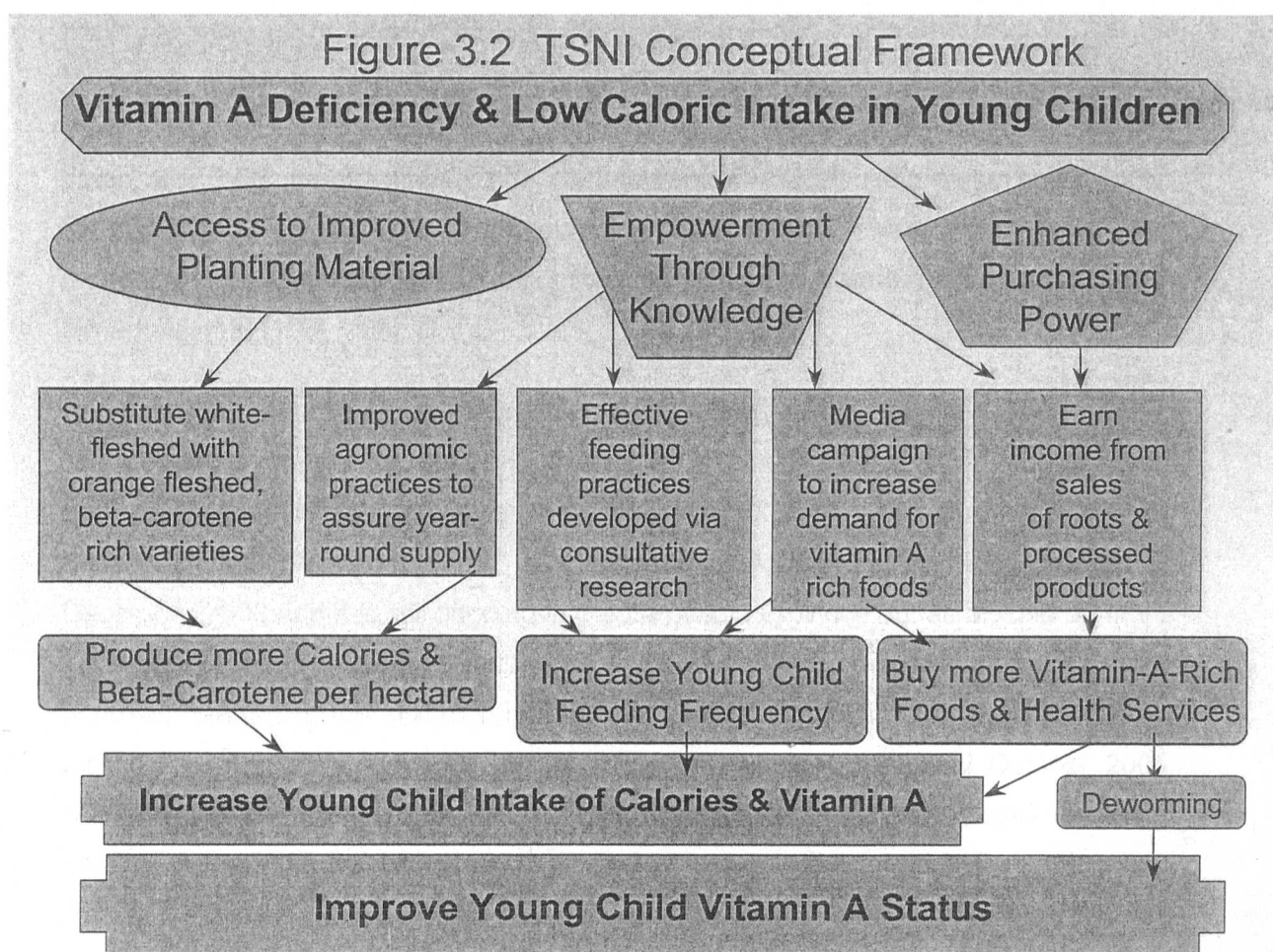
Zambézia Province

The intervention was based on a program model with 3 necessary and linked elements. To succeed in substantially increasing production of OFSP and vitamin A intake, the project aimed to simultaneously: 1) increase farmers' access to OFSP vines; 2) increase nutrition knowledge and create demand for OFSP; and 3) ensure sustainability through market development. [16] [105]

71

3.4). Participants had the opportunity to attend from 9 to 12 sessions (depending on site) over a 1-y period. The communication strategy also included radio programs, community theatre, prizes of promotional hats and *capulanas* (i.e. lengths of cloth worn as skirts), and market-based advertising to create an enabling environment for modifying behaviours and to create demand for OFSP and other vitamin A-rich foods. Along with the market-based advertisement for OFSP, the project also created an infrastructure for the supply for this cultivar, not only at the community level but at the district level, as well as in the capital of the province. The idea here was to allow the study communities to generate income while at the same time increasing demand and making it easier for people to have access to this cultivar. The control areas did not receive any of these activities. However, there was no way to avoid their exposure to the radio program since this was broadcasted to the whole Province. [16] [105]

Figure 3.2. Depicts the project's conceptual framework, where one can see three distinct elements working in synergism.

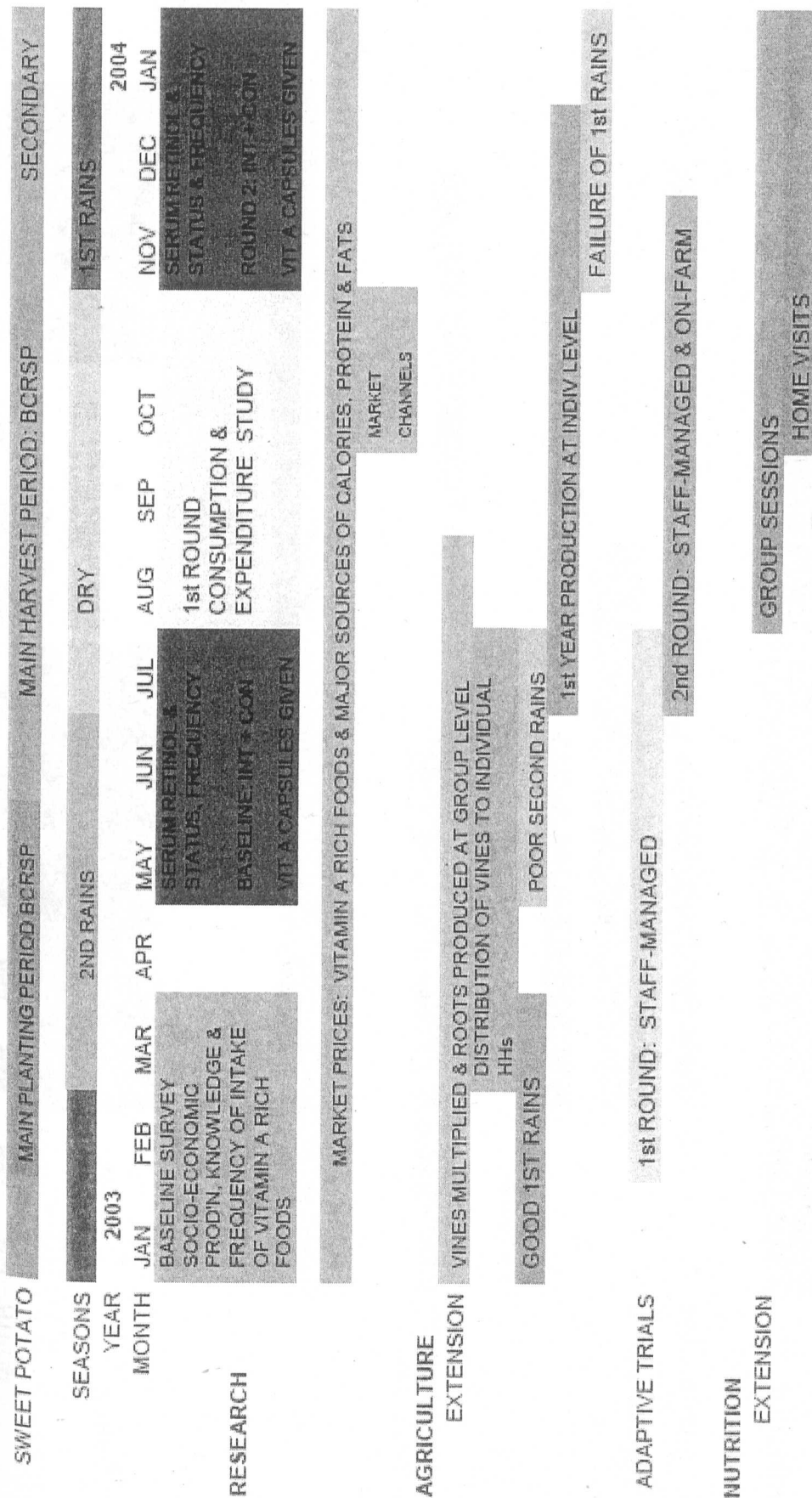


Apart from the program activities, the project also carried out repetitive data collection activities that had to be very well integrated with the former. Nine surveys were undertaken which gathered information on socio-economic and demographic characteristics of households (two rounds: baseline and end of study), agricultural production (two rounds: baseline and end of study), food and non-food expenditures (two rounds), child morbidity (which was based on the recall of the principal caregiver) (four rounds including baseline and end of study), adult and child anthropometry (four rounds including baseline and end of study), parental nutrition knowledge (two rounds: baseline and end of study), food frequency intake, dietary intake, and biochemical indicators. The aims were to: 1) establish baseline comparability; 2) characterize intermediate changes in the pathway toward impact (e.g. knowledge, OFSP production and consumption, vitamin A intakes); and 3) measure impact on serum retinol concentrations. Most topics were covered in multiple surveys to capture change across the life of the project. Dietary intake or (the consumption survey) was assessed using 24-h recalls (both at household- and individual reference child level) during the sweet potato harvest season each year. A simple FFQ determined how many days in the last week the child had eaten OFSP and other vitamin A-rich foods and fat sources; the FFQ was repeated 7 times to capture seasonal variation and change across the 2 years. The key outcome variable for vitamin A status, serum retinol, was measured four times at 5.5-6.5 month intervals for all intervention children and three times for all control children. In addition to the survey data collection, sweet potato plots were measured annually and prices of major sources of vitamin A were monitored monthly in 5 markets. [16] [105] [106]

The following figure 3.3 captures both the timeline of extension activities (i.e. agriculture, and nutrition) as well as data collection activities. As can be seen from the table, activities were occurring simultaneously. Agriculture activities started in December 2002 and it is still ongoing in the Zambezia Province but as a scale up of the TSNI study (which was only a Pilot). The nutrition education activities using the group approach started at the end of July 2003 and finished approximately in July 2004 for most women' groups, however for some groups, classes went until October 2004. Nutrition education in the form of home counselling, started around September 2003 and finished at the same time as the group classes. Measurements for each output step in the current study were taken within the group and individual nutrition education activities that occurred from April-May 2004 until October 2004. Attitude measurement were taken after the nutrition education activities had finished and measured impact of these.

FIGURE 3.3

TIMING OF RESEARCH AND EXTENSION ACTIVITIES: 2003-2005



RESEARCH

ASSESSMENT OF PRODUCTION 2003 & FREQUENCY

SERUM RETINOL & STATUS & FREQUENCY

ROUND 3: INT ONLY PLACEBOS GIVEN

2ND ROUND CONSUMPTION & EXPENDITURE STUDY

SERUM RETINOL & STATUS & FREQUENCY

ROUND 4: INT + CON

VIT A CAPSULES GIVEN

FINAL ROUND PRODUCTION 2004, KNOWLEDGE

MARKET PRICES: VITAMIN A RICH FOODS & MAJOR SOURCES OF CALORIES, PROTEIN & FATS

AGRICULTURE
EXTENSION

VINES MULTIPLIED & DISTRIBUTED TO INDIV HHs
2nd YEAR PRODUCTION AT INDIVIDUAL LEVEL

LOUSY 1ST RAINS

EXCELLENT 2nd RAINS

3rd ROUND: STAFF MANAGED & ON-FARM

ADAPTIVE TRIALS
MARKETING

COMMERCIALIZATION & MARKET DEVELOPMENT PILOT INITIATIVE

NUTRITION

GROUP SESSIONS
HOME VISITS
PROVINCIAL LEVEL RADIO PROGRAMS

Output steps-
measurement
in the current
study

Attention obs and recall and comprehension lexical meaning (inc. pilot)

Skills + recall + comprehension
psychological meaning (inc. pilot)

Attitude

3.3 Characteristics of TSNi study population – the mothers

All participants in the TSNi nutrition education program were the biological mothers of the TSNi reference children. The exception being for some women who though enrolled in the TSNi agricultural component, were not part of the TSNi research study (as they did not have any children under 32 months of age at the recruitment period).

The following is a list of characteristics (taken from the baseline survey) of the 560 (intervention groups only) mothers of the TSNi reference children:

- Age: varied from 15 to 56 years old, but average was 27.44;
- Education: 61.4% no formal education and only 0.5% had 7 years of schooling;
- 72.1% spoke no Portuguese;
- Marital status: 14.8% officially married, 74.1% in a marital union, 4.8% single, 6.8% divorced or separated, 2% widows, and 0.2% polygamy;
- 10.2% of the mothers were the head of the household, 82.9% were the spouse of the head of the HH;
- Agriculture principal activity of 99.1% of the mothers;
- 74.5% of the mothers actively participated in religious activities;
- Mother's nutritional status: average BMI was 20.30kg/m² but 17.9% had a BMI <18.5 kg/m²; average Hb level was 11 g/dl but 15.7% had a Hb level below 10 g/dl;
- Average age for having the first child was 18 years old;
- > 45% of the mothers consulted their own mothers for advice about the diet of their children;
- 73.9% of the mothers thought that the reference child was growing up well;

3.4 TSNI Socio-Demographic factors (or exposures)

As discussed by Ajzen (1992)[107]: “Age, gender, intelligence, self-esteem and other individual differences among receivers are rarely found to have strong effects on persuasion, and the results of different investigations are often inconsistent”.[107] However, Ajzen discussion is based on specific empirical studies carried out in the 60s and in a developed country context which is very different from a developing country one.

For the current study, throughout the results chapter, the output steps (i.e. the outcomes measured for the current study) will be associated with other factors measured during the TSNI baseline survey. These were factors that, on the basis of hypothesis and other past research[78], were expected to affect the output steps. These exposure factors are: age of the mother of the reference child (i.e. the study subject), her education level, the education level of the father of the reference child, household size, income per capita, and the mother's perception of the health of her child. The hypothesis in the current study are: that an older age might lead to the mothers having more schooling years and interest in the classes; having more schooling years might mean that the mothers might perform better in the output steps; more schooling years for the father of the reference child might mean that he might persuade the mother to have an interest in the nutrition classes; a bigger household size might mean that the mother has less time and therefore cannot attend so many classes; and a better income per capita might mean that mothers are more open to make changes.

Mother's perception of her child health status was measured in order to see if mothers that perceived her child to be in poor health would have higher attendance levels to nutrition education classes. The hypothesis is that they would, as they are the most concerned with their children health status and hence are aware that they could use the information gained in the classes to improve the health status of their ill children. This indicator was used instead of anthropometric measurements simply because it was shown that there was conflict between the mother's perception and the researcher's (i.e. mother perceived the child to be healthy when in fact the child had low nutritional status as measured by the nutritional status).

The goal is to explore and not explain, the variables that might contribute most to explain each of this study's output steps.

The following is a brief account of how they were measured:

- *Age of the mother (in years)*: Collected in the baseline survey by asking the subject and certifying with other members of the family;
- *Education level of the mother and father*: The formal education system in Mozambique consists of 12 school grades which make up the primary and secondary education. Only the completed education level or grade of the mother and father was collected during the baseline survey;
- *Household size*: It was also collected in the baseline survey and included all those subjects that usually resided and had their meals in that household. This was only associated with the exposure output step;
- *Income per capita*: An income proxy was developed based on aggregated data of salaried and self employment activities and sales of agricultural activity's products. To this total income, any costs of production were deducted;
- *Mother's perception of the health of the child*: instead of using nutrition education indicators such as *Height-for-age*, *weight-for-age* and *weight-for-height*, it was thought that the mothers perception of the child's health would be a better indicator, the following question *Do you think your child is growing up well and is healthy?* was part of the baseline survey. This was only associated with the exposure output step;

The sample population for this research was compared to the TSNI population for its representativeness on the above factors. Means were compared between the population measured for each output step and the TSNI sample population. An independent-Samples t-test was used for these comparisons, and groups were compared according to the intervention status (i.e. intervention I or intervention II). Results (not presented) of the comparisons showed that the current research sample population was representative of the TSNI study population for the majority of the socio-demographic factors. The few statistically significant differences between the current study sample and the entire TSNI population were small and may have resulted from multiple statistical testing and resultant Type one error.

3.5 Group class and individual class approaches in nutrition education – the TSNI case

In this research study, McGuire's model was used to measure and compare the levels of persuasiveness of the mothers exposed to the nutrition education program in either a group class setting (Intervention I) or in a group class setting plus an individual class setting (intervention II).

3.5.1 Input steps – the TSNI case

As was seen in section 2.2, one can manipulate input variables and their characteristics in a communication and measure their communication effectiveness by looking at the output steps. For example, as was seen in the literature review, a few studies were interested in seeing the impact of different kinds of message characteristics.

This research study is not looking at the influence of source, or message, or *receiver characteristics* on the persuasiveness of a communication but of the *channel* of communication, and more specifically at the mode of delivery. Nonetheless, the use of other input steps in the TSNI study will be briefly described:

Source: message was delivered by mid-level government nurses with 10 or more years of experience in government clinics who were trained to carry out field work related to diet, breast feeding and complementary feeding teaching, and to counsel mothers.

Each of these health promoters were then allocated to cover a locality and were set up to live on these localities during the time of the project.

Study participants were aware that the health promoters were government health staff, this being a positive aspect since nurses are trusted, are considered credible, and are valued for their health expertise.

Each health promoter worked in different ways with the women's groups. Some could have been friendlier, more committed, more motivated and more persuasive than others and that could have led to a higher degree of persuasion in some localities in comparison to others. However, these and other characteristics of the individual health promoter were not measured and therefore cannot be easily assessed. Nonetheless,

results for group class attendance levels showed that attendance was significantly higher in the locality of Malei compared to Posto Campo/Luala and Mexixine and non-significantly higher than Catale. Mean attention levels were non-significantly higher for Malei, but sample size was too small. Attitude scores were mostly higher for Malei. The health promoter in Malei was known for being the most enthusiastic and motivated of all the health promoters.

As health promoters taught in both group and individual classes, source should not be a factor influencing a difference between group and individual class level of persuasiveness.

Receiver: For both intervention groups, receivers were the mothers of under five children recruited to the TSNI study and registered into the women's group. TSNI baseline results showed that there was no difference in demographic characteristics and socio-economic status between the study population in both intervention groups.

Message: message delivery, like for example, the speed, loudness, clarity of speech, vividness and other aspects of speech was also influenced by the health promoter's style of communication.

Channel has an impact on the characteristics of the message. It will not be possible to know if message differences are the sole reason for a difference in the persuasiveness level between the two approaches or if it is the processes working within the approaches. Because message characteristics were different between the two approaches it will not be possible to discriminate between the what has what influence on the output steps. Instead, different message characteristics will be considered as intrinsic aspects of the two approaches.

The following is a description of the set up and activities of each approach and their message content.

Group class approach activities:

Group classes which started at the end of July 2003 and lasted a year were carried out with each women's group registered as part of the agricultural extension program. Table 3.1 provides a lay out of the several topics within the year of the nutrition education program. A detailed description of each topic follows.

Table 3.1 Topics given in the group classes

2003						
Month	July	August	September	October	November	December
Topics	Undernutrition in children	Food groups and Vitamin A	OFSP recipe	Puppets -- Complementary feeding	Enriched porridge	Story on complementary feeding
Tools	Photographs	Poster and photos of foods	Demonstration with participation from the women's groups	Puppet show	Recipe demonstration	Ministry of Health' flip chart
Content	To raise awareness of causes of malnutrition and how to identify it and prevent it	Eat from the 4 food groups	Enrich the porridges with local foods and OFSP	A role play describing the optimal diet for the child at different ages	Enrich current porridges (i.e. cassava flour) with locally available foods	Describing the optimal diet for the child at different ages
2004						
Month	January	February	March	April	May	June
Topics	Ana's pregnancy	Story of little John (part 1)	Story of little John (part 2)	Hygiene practices	Story of little Adelina	The 6 recommendations
Tools	Ministry of Health' flip chart	Photos and picture cards	Photos and picture cards	Picture cards	Ministry of Health' flip chart	Picture cards
Content	Care during pregnancy issues	Describing the optimal diet for the child at different ages and change taboos related to it.	Describing the optimal diet for the child at different ages and change taboos related to it.	Describing several hygiene practices to avoid diarrhoea	To raise awareness of causes of malnutrition and how to identify it and prevent it	Similar to 6 key recommendations given at counselling sessions
Overall main messages	Nutrition messages emphasized daily use of vitamin A-rich foods, appropriate breastfeeding and complementary feeding practices, and increasing young child feeding frequency so that the child will not get malnourished. Also, frequently weighing and taking the child for check ups to monitor the child' health status and growth					

Although men were encouraged to participate, the vast majority attending the nutrition sessions were women. Classes were lead by the health promoters that visited each farmer group every 3-4 weeks. Attendance to a topic was monitored so that in case of low attendance the topic was repeated. Classes were always given in the local language, Chuabo.

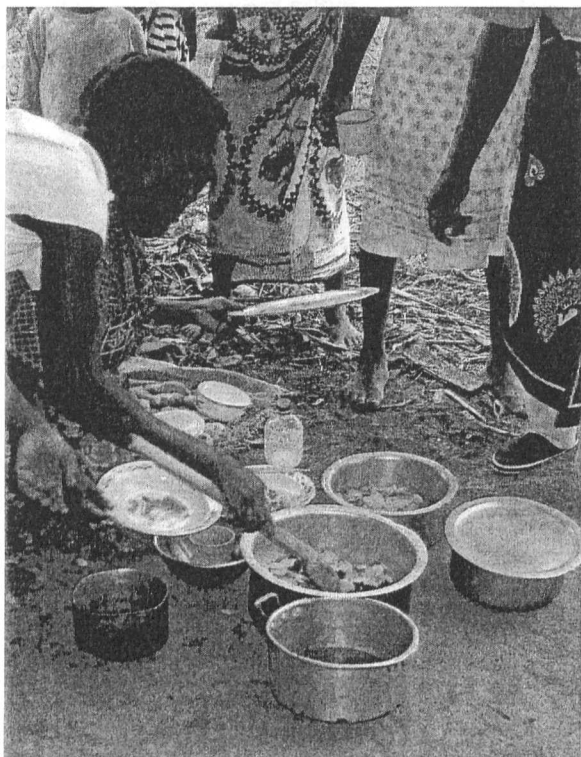
Group sessions covered the following topics:

1. *What is malnutrition?* Mothers were taught to identify the different types of malnutrition, their causes and how to prevent it. Photographs of children suffering from Marasmus and Kwashiorkor at different stages of the disease were shown to the mothers.

2. *The food groups and vitamin A.* Mothers were taught about the four food groups, the benefits of each group and what kinds of locally available foods can be found in each group. They were taught that it was important for the child to eat from the 4 food groups every day and meals and dishes were planned as an example of a balanced and varied diet. Mothers were also taught about vitamin A: What is it, what are the benefits, especially for the child, and what are the main sources of vitamin A. A poster and pictures of foods were shown to the mothers.

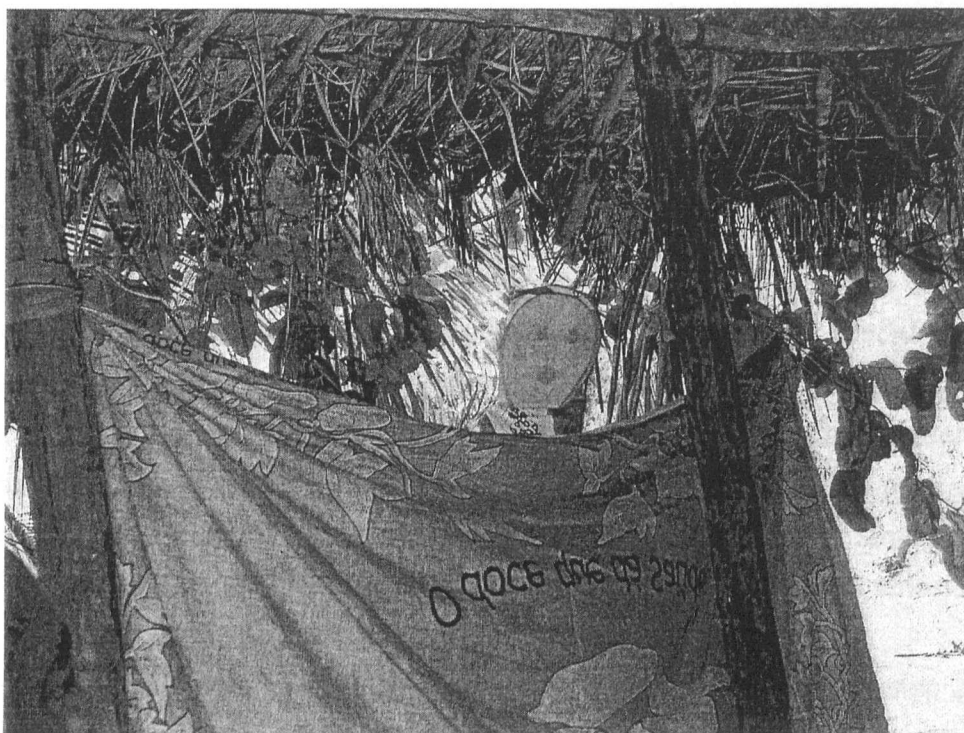
3. *Recipe demonstration using orange fleshed sweet potato (OFSP).* Participatory demonstrations of 3 or more recipes for weaning foods using OFSP and locally available foods, especially sources of protein and fat (see figure 3.4).

Figure 3.4. Recipe demonstrations with the women's groups



4. *Puppet show on complementary feeding* (see figure 3.5). This puppet show was a conversation between two mothers, one that had just come back from the TSNI nutrition classes on breastfeeding and was very motivated and the other had not attended any class and needed some convincing. It gave an opportunity to put what has been learned in perspective and adapted to the local costumes. The mother that had just come back from the class, was trying to show the other that there was a strong link between children's diet and high prevalence of illness and mortality. She also gives examples of two children in the village that suffered from kwashiorkor and marasmus and what their parents were feeding them. She then describes what was said in the class about breast feeding and complementary feeding until the age of 2. It gave an opportunity to talk about a lot of the taboos regarding colostrums and breast feeding and to explain again the recipes using OFSP and other locally available foods. It also included messages on feeding practices such as that the child should be encouraged to eat, should eat on his/her own plate and should have clean hands before eating. Two puppets were used and played by the health promoter behind a cloth ("capulana") advertising the OFSP.

Figure 3.5 Puppet show



5. *Story on complementary feeding.* A flip chart developed by the health ministry was used for this class. It emphasized the period where complementary feeding should be started (at the time was at 4-6 months) and that before that the child should only have breast milk. It described what should be the optimal diet for the child from the time that it should start complementary feeding until the age of 1. It described what should be the frequency of eating at different ages, and what to feed the child, including examples of the mixture of foods. It emphasized that the child should still be breastfed. It also included a message on hygiene, and another on the importance of taking the child to the health centre for growth monitoring.

6. *Recipe demonstration using locally available foods.* Participatory demonstrations of recipes for weaning foods using maize and cassava and emphasizing the importance of adding a locally available source of protein and fat.

7. *The diet and care of the pregnant mother.* A flip chart developed by the health ministry was used for this class. It emphasized the care that a pregnant woman should have when she is pregnant. It teaches mother how many times a day they should eat, what kinds of foods hey should have, that they should work less and rest more and go for frequent check ups at the health centre.

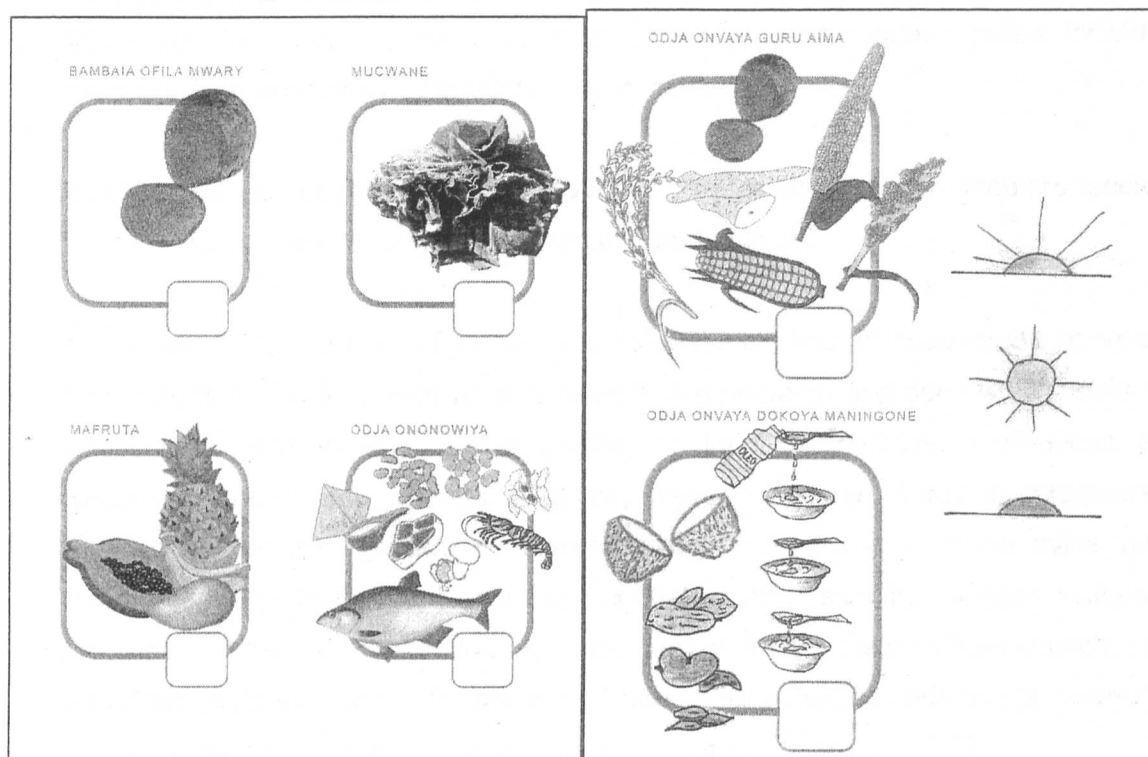
8 & 9. *The story of little John*. Picture cards developed by the TSNI were used for this class. Every message was given along with a picture card. Cards showed pictures of a specific child, little John, and how he was fed since he was born until he was five years old. The cards showed a picture of the child at different stages (very little and still breastfeeding, bigger but still not crawling, crawling, starting to walk, walking, and playing football) and at each of these stages there were recommendations on what to feed the child and the frequency of feeding. Key messages were emphasized. Moreover, practices like for example giving water, solid foods and special teas when the baby is born, interruption of breastfeeding when the mother is sick, pregnant and the child is sick were described as inappropriate. Other recommendations were given: boiling water for child, wash hands before feeding, wash plates and pans, feed a source of protein every day, give a source of fat 3 times a day, dark green leaves every day, child should eat on his own plate but should have supervision at all times.

10. *Hygiene practices*. Picture cards (taken from the PHAST step-by-step guide: A participatory approach for the control of diarrhoeal diseases, WHO) showing the modes of transmission, based on the F-diagram, were shown and explained to mothers.

11. *The story of little Adelina*. A flip chart developed by the health ministry was used for this class. The flip chart explained the conditions/situations that could lead a child to becoming malnourished, and what could be done to prevent it and to treat it. It also emphasized the need of taking the child frequently to the health centre for check ups.

12. *The 6 recommendations* (see figure 3.6). Mothers were given a card with drawings for 6 recommendations: give orange fleshed sweet potato to your child every day; give dark green leaves to your child every day; give fruit to your child every day; give a source of protein to your child every day; give a source of fat 3 times a day; give a source of carbohydrates 3 times a day. They were then asked to practice the recommendations at home and to counsel their neighbours with the same recommendations.

Figure 3.6. Handouts of the 6 recommendations given to mothers



Teachings of these topics usually lasted between 30-40 minutes, however topic's teaching were preceded by weighing sessions bringing the duration of the whole session up to an average of 1 hour. As can be seen, there was quite a lot of repetition of messages through out the topics apart from the fact that topics were usually repeated at least once more. Classes comprised of an average of 14 participants (ranging from 6 to 23). Classes were usually carried out in the community meeting places, for example, in an open space under a tree and were usually carried out in the afternoon as the health promoters were busy doing individual classes in the mornings. Not only the mother-child pair recruited to the TSNI study participated in the nutrition group classes, but other women that were recruited to the women group but not to the research study. Because attendance levels decreased during the rainy season, probably because mothers were busy with planting, prizes like cups, plates and spoons were given as an incentive for attending classes.

Individual class approach activities:

This could be a one-to-one session in a health clinic, usually called individual counselling, or in community projects - home visits.

Individual classes started in August 2003 and also lasted for a year. Mothers received in total 6 visits, every other month, during the study period.

An adapted trials of improved practices (TIPs) is the method adopted for this approach. TIPs constitute the core method in a consultative research approach which offers the potential to gain in-depth understanding of feeding practices, motivations and constraints to behaviour change. Typically, several visits are made to caregivers in selected households. Current nutritional practices are analysed in the initial visits. During follow-up visits, the researcher negotiates some specific changes in feeding practices that the caregiver follows for a set period of time. Researchers subsequently learn from mothers which practices work. This process identifies effective and practical behaviour changes which are acceptable and feasible for families.[108]

As part of that work, 12 key recommendations for promoting feasible changes in dietary, feeding and hygiene practices were developed specifically for study children (the youngest having attained at least 6-7 months of age when the messages were developed), summarized as follows:

1. *Give orange-flesh sweet potato to the child every day, preferably accompanied by a source of fat;* Mothers were taught of the benefits of vitamin A, then persuaded to feed some OFSP (half a 250ml cup of cooked and mashed OFSP) to the child every day and if possible to add a source of fat as it increases the absorption of vitamin A. They were also taught to recognize others sources of Vitamin A, especially for those occasions when OFSP was not available.
2. *Give a source of carbohydrate-rich foods such as rice, wheat, maize, sweet potato, manioc, rhizome, sorghum, or millet, in each of the three main meals: breakfast, lunch, and dinner;* Mothers were encouraged to give one of the above foods on the 3 main meals as they were needed "for the child to have a lot of strength".
3. *Give green leaves every day to your child;* Mothers were encouraged to prepare green leaves every day for the child because they belonged to that group of foods that "protects the child from diseases".

4. *Give fruit to your child at least once a day and preferably in between the main meals*; Mothers were explained that the child should eat in between the main meals and that the mother should give ripe fruits to the child. The health promoter made sure that the mother understood what she meant by fruits and what kinds of fruits she could find locally. She explained that fruits were good because they belonged to that group of foods that “protects the child from diseases”.

5. *Put a tablespoon of fat in the three main meals of the child*; Mothers were taught which sources of fat were available locally and then encouraged to add a tablespoon to the meals of the child (if they did not already contain fat) in the 3 main meals as they belonged to that group of foods that “gave a lot of energy to the child”.

6. *Give at least one source of protein every day to the child*; Mothers were taught that they belonged to that group of foods that “helped the child grow well”, and exemplified by explaining that these foods will help a child grow well, just like a brick house needs good bricks for it to grow. Health promoters also explained that there are two types. That animal sources (and they went through the list) were the best source but that beans, nuts and seeds were also a good source, but that they would need to be mixed with rice or maize.

7. *Do not skip any of the three main meals*; Mothers were encouraged to always give at least 3 meals a day to the child, and that they should not skip breakfast (which usually happens often) because the child needs to start the day with a lot of strength and energy.

8. *The child should eat from his/her own plate*; Mother's were encouraged to let the child eat on his own plate, because when he/she shares with other older siblings the tendency is for the latter to eat his/her part. Also, it gives an opportunity for the mother to control the quantity that the child has eaten.

9. *The mother should assist her child in eating*; Mother's were encouraged to feed their children and not let them eat on their own as they are sometimes lazy to eat.

10. *The mother should stimulate the child to eat when he/she refuses*; Methods of stimulation were taught to the mothers, and they were asked to try them out when the child refuses to eat or leaves a lot of food in the plate.

11. *Wash hands before meals.* Mother's were taught the causes of diarrhoea, and were explained that both the child's hands and hers should be washed before meals and after the child or she defecates. She was encouraged to use soap, but in its absence, to use ash or clean sand;

12. *Boil the drinking water for the child;* Mother was also encouraged to boil water for the child as the child's body was not prepared for the water that adults drink.

The above 12 recommendations were repeated in each visit, therefore, 6 times overall. As can be seen, there were no recommendations on breastfeeding. All recommendations were carried out in the local language, Chuabo. Each recommendation was aided by photographs of foods, or drawings or poster (i.e. the food group poster).

The adoption/implementation of these 12 recommendations were closely monitored at each follow-up visit. During follow-up visits, messages were reinforced based on the 24hr recall of the child's consumption collected from the mother. The health promoter reviewed compliance to the 12 recommendations, congratulating the mother for her achievements, noted down any constraints to adoption and worked with the mother to seek a solution. Two out of the six visits were carried out during lunch time and consisted mainly of a structured observation of the meal time, recording meal preparation, what was eaten by the child, feeding practices, and hygiene practices. This gave an opportunity to identify problems which were then discussed during the individual class which followed the observation.

Both group and individual class approaches used verbal forms of communications and a mixture of mediums of delivery but both used a lot of visual aid. No session occurred without a visual aid. Group classes used flip charts, picture cards and posters and individual classes used picture cards, posters and drawings. Apart from this, group classes also used puppets, role plays and demonstrations.

Looking at the message characteristics used in both approaches it can be concluded that even though both used a lot of repetition, this was much higher and more systematic in the individual classes. Moreover, messages in individual classes were in fact key messages, much shorter and straight to the point, while in the group classes, messages were incorporated into a story. In both approaches, messages tended to follow a pattern: explaining what is required of the mother and why is it beneficial for the child (i.e. promise appeals). This is true even more so in individual classes. In

group classes, some messages also contained fear appeal. In both approaches the language in the messages was very simple and clear and did not contain any technical terms, and if it did (i.e. malnutrition, vitamin A), the health promoters would explain it. In both approaches humour was hardly used. In group classes, many messages would acknowledge and refute opposing arguments. For example, all the practices that were known to be incorrect by international standards, were mentioned, followed by an explanation of why mothers should not practice them, appealing to explanations of why they were bad for the child.

Organizational features of the individual class communication

In a book written by Ley [109], he discusses the best ways of providing health related information to patients and clients. He does this in the context of the communication process, looking at the so often problems encountered by the patients when receiving a communication from health care providers: reduced satisfaction with the communication, poor understanding of what they are told, reduced memorisation and recall of what they were told, low compliance to instructions – and at what can be done to improve all these problems in the communication. For example, in order to increase recall and understanding he proposes that:

- There should not be too many statements as the more statements there is (or information presented), the harder it is to recall them,
- Information/advice should be simplified whereby shorter words and shorter sentences are used, making the information/advice easier to understand;
- Information/advice should be repeated by the communication or the receiver;
- Information/advice should be specific rather than general. For example, instead of saying that the child should eat more meals a day, one should say that the child should eat 3 main meals a day, in the morning, at lunch time and dinner time.
- Should use explicit categorisation which is a technique whereby the communicator tells the patient what categories of the information are to be provided. For example: “ Now I am going to tell you: what is wrong with you, what tests will be needed, and what the treatment might be.” Compared to a non explicit organization, an explicit organization increases recall;
- The most important messages should come first since people recall best what they are told first;

Even though not mentioned in the above reference, another feature that should also be included is a summary of the main points, especially if the information is a lot or maybe difficult.

Using the above features as a guide for the optimum development of the messages given at the individual classes (i.e. individual counselling in the TSNI) it can be discussed that perhaps 12 key messages was a lot to memorise, especially for that population which was illiterate meaning that a reminder could not be used since they could not read it. Maybe 6 messages were more appropriate.

The information for each of the 12 key messages was pre tested for its ambiguity and clarity and it was observed that the message was easily understood by the mothers. One of the reasons for this, as suggested by Ley [109], is that short sentences were used for each key message, as well as visual aids.

Also, as suggested by Ley, the key messages were repeated in 6 different occasions (i.e. every 2 months mothers would receive a visit from the health promoter who would repeat again the same 12 recommendations.

Again, as suggested by Ley, the recommendation was specific and not general. For example, mothers were told to add 1 tablespoon of an available source of fat to the child's meal on the 3 main meals: breakfast, lunch and dinner.

The technique of explicit categorisation was not used, but then Ley suggests it use when the medical information given was extremely complex which was not the case with these messages.

As described by Ley, it is usually the first messages that are recalled best. Hence, the most important message, the one most related to the TSNI objectives, was the first recommendation given to the mothers – “give OFSP to your child every day”

Since messages were already quite summarised, a summary of the messages was not given to the mothers.

Moreover, all messages were followed by a visual aid showing an example of the action recommended. As well to aid understanding and recall, messages were organised in categories of recommendations related to the food diversity, frequency of eating, feeding mechanisms and hygiene practices.

3.5.2 Channel – group and individual class processes and the TSNi case

In the literature review (Section 1.5.1), programmes on nutrition education in developing countries have either used mass communication or face-to-face communication, specifically the group and the one-on-one approach, as the sole modes of delivery of the message. The TSNi program used both types. It used radio programs and painted murals for the mass communication program. Mass communication has the advantage of targeting a wider audience but not necessarily the target population. In the case of the TSNi study, illiterate mothers were not able to read murals or posters, and even if they could read, most of them did not frequent that market place. Also, only 1/3 of the households in the TSNi study possessed radios. Mass communication was not therefore the best approach for nutrition education to the TSNi research study target population.

3.5.2.1 Group class approach

The group class approach used in TSNi should not be confused with group counselling which generally involves a relatively smaller group of maybe 5-10 subjects. Group counselling is an interactive form of education in which learners generate the specific topics to be addressed and share their knowledge and experience with other group members through discussion. The educator becomes a facilitator who, rather than lecturing, encourages participants to participate in discussions and to discuss freely among themselves their own approach to the nutrition problems posed during the session and comes in only to correct misinformation and manage group dynamics. It is an approach that allows the exploration of values and beliefs and deal with sensitive issues.[110]

In contrast, the group classes used in TSNi involved a health promoter giving a class to a group of participants, very similar to a class in a primary school (see figure 3.7). These were relatively less interactive and more top-down and didactic than group counselling, hence, involving little group processes. Also, the groups were much bigger, and sometimes it might even have reached 40 participants. Most of the time, there is a lot of noise coming from passer-by and other members of the village, but in the case of the TSNi study, the majority of the disturbance came from the small children in the participants laps who every so often would cry and usually required the mother to stand up to attend to the child. The participants are not comfortably seated as in the majority of times they have to sit on the floor.

The group class approach is more of an information giving kind of approach where the health promoter interacts with the audience by asking herself questions and asking them if they have questions. If some participants had some questions, then a small discussion might have taken place, but the experience in the TSNI has shown that the majority of mothers were quite shy and usually only a few ask questions. Men seemed to contribute the most in the classes, whenever they were present. However, their presence in the classes seemed to hinder the women's participation.

The health promoter was usually more concerned with giving the information and hence, could not give individualized attention to each of the participants. This meant that it was difficult for her to notice if participants were inattentive and if they were agreeing with the messages.

Even so, the advantage of the group class approach is that there is more peer contribution and support, so people can relate to each other, and in a country like Mozambique where taboos are strongly rooted in the community and the only way to change them is by changing the beliefs of the community, then group classes might be a better method of changing beliefs and attitudes. A positive thing in the group class program was the inclusion of participatory demonstrations which not only lead to more interaction between participants and between participants and the health promoter, but it also allowed for skills-gaining and "empowerment" of mothers.

It would be possible to transform the group class into group counselling, but the number of participants would have to decrease and hence, more sessions would have been needed to cover the whole study population. But, the main barrier is that the facilitator would have to be someone highly trained, which is difficult for developing countries as not only is it difficult to find qualified staff but very costly.

Figure 3.7. Group class



3.5.2.2 Individual class approach

The individual class approach involved only two people, i.e. one health promoter educating one subject (mother) on nutritional matters (see figure 3.8). These meetings usually occurred in the mother's home, which made it more convenient for the mother. Also, the mother was more comfortable at home where there is not so much noise and disturbance. Moreover, this approach means that the health promoter can dedicate more time and attention to just one person, making sure that the mother understands what is being taught. The health promoter has the opportunity to read in the mother's expression either agreement/disagreement, boredom, confusion, etc, which gives an opportunity for intervention if necessary. Also, the health promoter has the opportunity of increasing attentiveness if she notices that the mother is becoming inattentive, by for example pausing for a while.

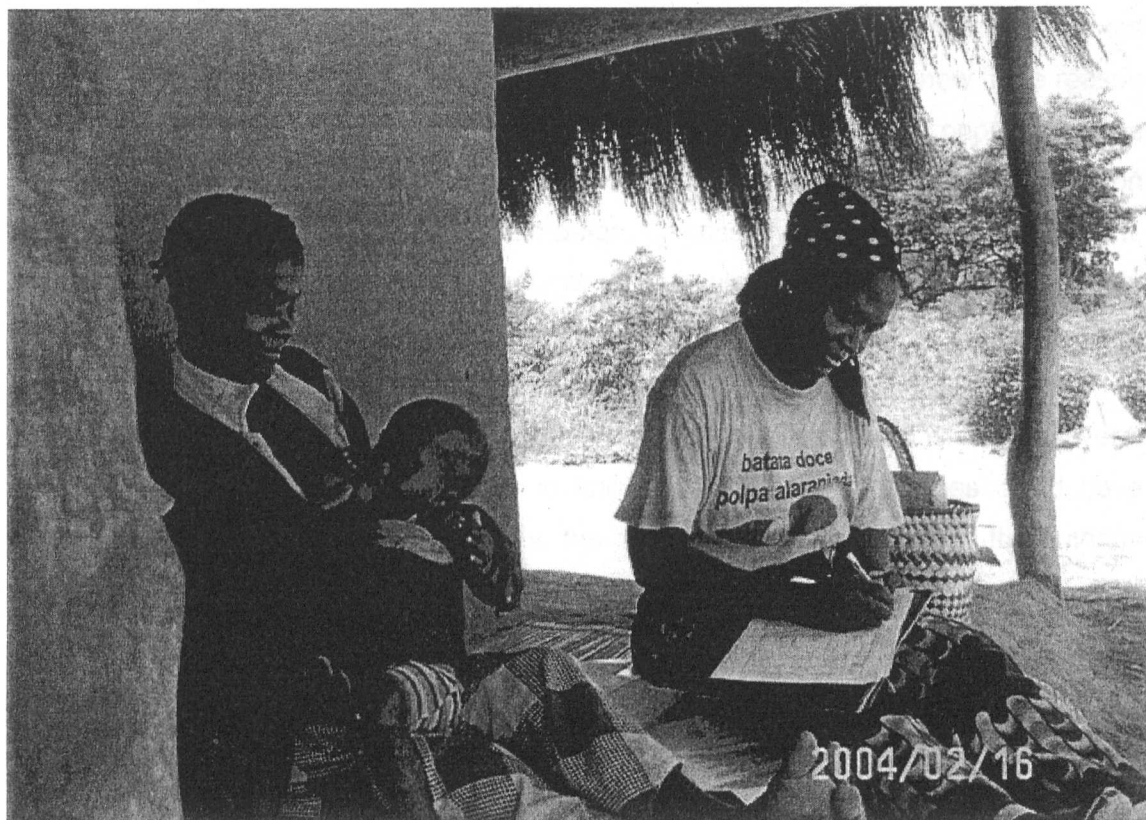
Education is also more collaborative than in a group class. Due to the method used for this approach, TIPS, this was a much more persuasive and tailored education, where the health promotion tried to convince the mother to follow a recommendation, and discussed with the mother how to carry this out taking into account the particular situation of that mother. There was also much more opportunity for them to become familiarised with each other, since mothers were visited 6 times.

In the TSNI study, mothers were actually more shy in the one-to-one approach than in group classes. The tendency was for them not to look directly to the health promoter during the counselling, but usually look down at her lap or at the visual aid presented by the health promoter. Moreover, just like in the group classes, the presence of the husband would make the mother even more quiet.

The biggest disadvantage of this one-to-one approach is that it is much more time-consuming and more costly than the group class approach. While with the one-to-one approach only one mother is exposed to the nutrition communication, with the group class up to 20 mothers could be exposed to the nutrition communication. It is also argued that repetition of messages aids memorisation, recall and attitude change, but also adding new arguments has some persuasive advantages over simply repeating old arguments with regard to retaining attention.[71] So it is possible that the constant repetition of the messages in every visit led to some boredom on the part of the mother.

Taking the example of the TSNI study, what should have been a counselling session, became more of an information giving session, with repetition of the messages and some discussion about the reasons why the mother did not manage to implement the recommendation(s) and some negotiation for the mother to implement it. However, it was too automatic and most of the times, the health promoter did not have the skills to analyse the situation of that particular household which was needed for an effective persuasion and negotiation. They also lacked improvisation skills and initiative.

Figure 3.8. Individual class session



This section has described and discussed the difficulties within the two most used face-to-face nutrition education approaches in rural settings of developing countries. From these descriptions it is possible to conclude that the one-to-one approach has more advantages regarding the processes within the approach in comparison to the group class approach. However, will these advantages make a difference and hence make the group class + individual class delivery method score significantly higher than the GC only delivery method?

3.6 Complex interventions

The TSNI intervention can be considered as a “complex intervention” since it comprises of a number of separate elements which working independently or inter-dependently, seem essential to the proper functioning of the intervention. Complex interventions are health service interventions that are not drugs or surgical procedures, but have many potential “active ingredients.” The greater the difficulty in defining precisely what exactly are the active ingredients of an intervention and how they relate

to each other, the greater the likelihood that you are dealing with a complex intervention.[111, 112] In the TSNi case, it integrated an agronomic, a nutrition and market development component. All of this interchangeably make it difficult to pick out the intervention (s) that had a clear impact. For example, we do not know if the reason why there was an increased consumption of OFSP by the TSNi study population, was due to the agriculture extension activities alone, or the nutrition education intervention alone, or maybe the social marketing campaign. Or possibly a mixture of some of them, or all of them. Hence, the many elements of the intervention have complicated or biased the outcomes of the intervention.

To make it even more complex, the nutrition education intervention consisted of several different messages, making it difficult to know which of the messages could have contributed to certain behaviours. Briefly, the main messages attached to the nutrition education intervention were:

- Colostrum is essential for the infant
- Breastfeed exclusively until the 6 –month
- A pregnant or sick mother can still breastfeed. It is very important to continue breastfeeding the child if he gets sick
- For those children already weaning, it is important that he/she eats frequently a diversified diet, otherwise the child might become malnourished
- The mother should weigh the child frequently to prevent the child becoming malnourished
- The mother should feed at least 3 times a day: morning, lunch time, and dinner time
- The mother should feed the child everyday foods from the 4 food groups
- Give OFSP every day to the child and other vitamin A rich foods.

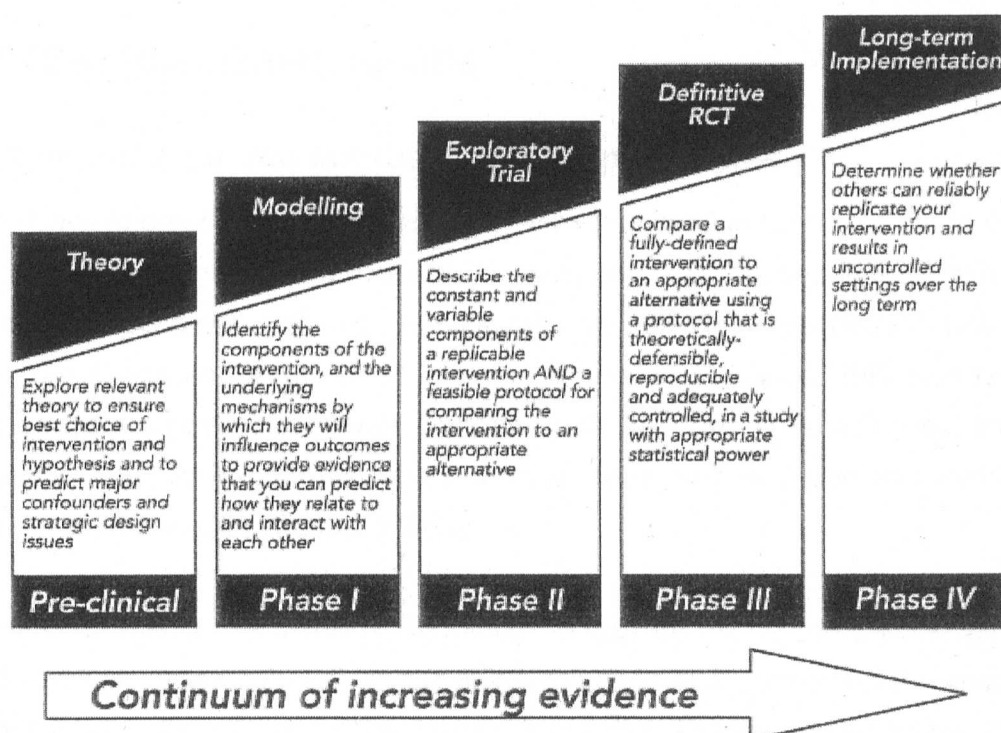
The above messages covered issues related to recommended breastfeeding practices, causes and prevention of malnutrition, including recommended complementary feeding practices, as well as messages encouraging the consumption of locally available vitamin-A rich foods, especially OFSP. However, as can be seen from the description of the various topics taught in the group classes, the nutrition education program also included messages on good care during pregnancy, good hygiene practices, and countering myths and beliefs that pose as barriers to good practices. While in fact the main emphasis of the TSNi study was on the promotion of vitamin A rich foods and increase in caloric intake. Hence, the breadth of the nutrition education itself was so extensive, covering a lot of areas that it is difficult to know which messages have contributed to which behaviours. The same for the individual classes, mothers received

not only 12 key messages but also in 6 repetitions. For the individual classes for example, it is difficult to know as it is if it is the content of the communication, or maybe the frequency of the communication, or a mixture of both, that was the “active ingredient”. At this level, this array of messages might interact with each other, complicate and bias the outcomes of the intervention.

Because of the above problem, MRC has developed a framework for the design and evaluation of complex interventions, which is described below in figure 3.9. [111, 112]

The MRC distinguishes 5 phases in the framework: 1) theoretical; 2) modelling; 3) exploratory trial; 4) main trial; and 5) long term surveillance.

Figure 3.9 MRC Framework model for the evaluation of complex interventions



The MRC framework suggests that investigators should carry out a literature review of the proposed area of investigation in order to high-light important issues that might clarify the complex intervention to be studied, help tease out the likely facilitating factors or barriers in developing the intervention. [111, 112] For example, if planning a complex intervention to reduce girls' vulnerability to HIV/AIDS infection, one should know the theoretical background on what factors make girls vulnerable, and one what has worked in making them less vulnerable and barriers to this objective. After looking at the literature, one should identify the key components of the intervention. At this stage it might be necessary to collect more information, through for example, focus

group discussions to understand the issues into a greater depth. Qualitative research might be helpful in identifying which are the “active ingredients” of the complex intervention, and which elements are not related to the treatment effect. For example, maybe one of the main causes of girls’ vulnerability is high drop out and low enrolment to schools, hence, the intervention has to include a component that would improve this factor, but perhaps counselling from health services has no impact. The MRC also suggest to map out the intervention, by doing a diagram, or a flowchart, of the various components, which may be useful to identify the key relations among components and potential weak points in the intervention. Also, by doing a diagram one can start thinking about how and which components should be controlled and which should be varied. [111, 112]

3.7 TSNI (published) results

3.7.1 impact from the nutrition education activities

Attendance to group classes: Mothers who were supposed to attend the group nutrition education classes attended a mean of 7.8 sessions (out of 12 sessions). When asked for the main reasons for missing sessions, most women (87%) reported this was due to their own illness and/or the need to care for another ill person.[106] However, the biggest drop in attendance levels coincided with the rainy season, hence, the high probability that mothers did not attend because they were occupied with work in the field (i.e. their own agriculture activities).

Nutrition knowledge: At baseline, knowledge scores on the 12-point scale were very similar (mean scores 3.3--3.4) across all groups (men and women in intervention and control areas). At the end of the study, intervention women scored higher than at baseline (mean score 8.1) and higher than control women (mean score 4.3). Intervention men also improved their scores (mean score 6.3) and scored higher than control men (mean score 4.7) ($p=0.001$ for all comparisons). [106]

Dietary practices:

During the second 24-hour recall, higher proportions of intervention than control children in all age groups were fed three main meals a day. The largest difference was observed among the youngest children: at 6 to 11 months, 49% of children residing in intervention households compared with 17% of children in control households were fed during all three main meals ($p < 0.001$ for all comparisons). [113]

Also, diet diversity was higher among intervention than control children, with 32% of intervention children consuming food from more than four groups, compared with 9% of control children ($p < 0.001$). This more diverse diet derived from high consumption levels of foods from three groups: vitamin A-rich roots, vitamin A-rich fruits and vegetables, and legumes and nuts (principally groundnuts) [113]

3.7.2 Impact from the agriculture activities

OFSP production: Five of 9 introduced OFSP varieties were accepted by farmers both in terms of taste and agronomic performance. Households in farmers' groups expanded the area under production in y 2, increasing mean plot sizes from 33 to 359 m². 90% of intervention households were growing OFSP in the final year against 11% for the control households. Median total sweet potato production for intervention households increased from 73 kg at baseline to 127 kg during the final study year, of which 71% was OFSP, while in control households, median production remained very similar to baseline. The percentage of intervention households selling sweet potatoes increased from 13 to 30%, and in 2004 OFSP was the cheapest source of vitamin A on the market, costing 1 cent for 700 RAE. Agriculture projects that aim at increasing the supply of food in order to increase consumption have to be aware that the study participants might decide to sell their production instead of keeping it for the household consumption. This was also a key challenge for the TSNi study. The attempted solution was the market development strategy which encouraged households to produce surplus production for sale, while assuring sufficient home consumption. In year 2, the project introduced the concept of quality grades, whereby only those OFSPs that offered a certain quality (i.e. certain size/weight and no pest) could be sold. [113]

OFSP consumption: The increase in OFSP production also resulted in an increase in its consumption by the study children in both intervention and control areas, as was reported by repeated Food Frequency Questionnaires (FFQ). While changes between the initial and last FFQ were significant within each group ($p < 0.01$), the size of the increase was much larger in the intervention area. In the last 2 FFQ survey rounds, >50% of the children in the intervention area ate OFSP on at least 3 out of the last 7 days, as compared with 4–8% in the control area ($p < 0.001$). [106, 113]

Intake of vitamin A: This consequently resulted in an increase in the median intake of vitamin A which was much higher ($p < 0.001$) for intervention children (426 mg RAE) compared with control children (56 mg RAE). Intakes were also higher for the

intervention group for a range of other micronutrients and for energy. OFSP contributed 35% to the vitamin A intakes of all children in the intervention area and 90% in those who consumed it the previous day. [106, 113]

Serum retinol: The increased recorded intake in vitamin A lead to an increase in mean serum retinol in the study children. At baseline, mean serum retinol between intervention and control children did not differ either in all children or in healthy children (i.e. children with CRP values <5 mg/L). However, for the final blood sample collection, the unadjusted mean serum retinol concentration was higher ($p<0.01$) for intervention (0.74 $\mu\text{mol/L}$) compared with control children (0.67 $\mu\text{mol/L}$). Also, the between-group difference was larger (0.09 $\mu\text{mol/L}$) for healthy intervention children (0.80 $\mu\text{mol/L}$) compared with healthy control children (0.71 $\mu\text{mol/L}$; $p<0.001$). Moreover, by the end of the study, prevalence of low serum retinol was 10% lower ($p<0.01$) in intervention children compared with control children. In the healthy children in the intervention area, prevalence of low serum retinol dropped from 60% at baseline to 38%. In contrast, prevalence remained the same in healthy control children (52% at baseline vs. 53% at final blood sample collection) for a between-group difference of 15% ($p<0.01$). [106, 113]

CHAPTER 4

METHODS AND TOOL DEVELOPMENT FOR THIS STUDY

Chapter 2 explained that McGuire's model was the chosen model to measure persuasiveness levels between the two TSNI nutrition education interventions, and in Chapter 3 the study background was described. This chapter provides detailed information on the tools designed to measure the output side of the communication/persuasion matrix. McGuire himself has given a limited description on how to measure these output steps. Hence, this chapter's biggest contribution is its attempt to develop tools for assessing the various output steps. This might lead to problems with respect to the development of precise assessment tools. In part, because of their novelty: some of the methodologies for evaluating the various output steps within educational settings and, specifically, within rural developing country settings have not been developed. As well, there is an issue of transferability of this model, which was developed and used to guide interventions in developed countries, to a different social and economic context. Issues that might be important to take into consideration are: different culture, different beliefs, resource-constrained environment, different social context, and the high illiterate level. Regarding the latter, it seems even more important to measure the output steps in this context since for example, illiterate levels which are much higher might have a higher negative impact on the learning process of the mothers.

As a response to the above situation of transferability, the tools developed in the present research study have been modified and adapted from tools used in research in cognitive psychology, mostly in laboratory settings, but also from studies carried out under natural conditions. Nonetheless, compared to the studies mentioned in the systematic review which used predominantly subjective evaluative methods of assessing the output steps, this study conducted objective evaluations and additionally conducted direct systematic observations of some output steps.

The current research study is among the few studies to have used McGuire's output steps to evaluate a nutrition education program, but the only one to present results in a developing country setting. Due to constraints in the experimental design, the current research used a simplified 6- output step and some what modified version of McGuire's model. Interest/Liking and Decision Making were not measured, as the self-reports

often used to measure them (i.e. like for example asking participants to rate how much they liked a communication or asking them to rate their intention to perform the behaviour recommended by the communication) were not considered appropriate for the current research. Even though behaviour change (as change in dietary practices) was measured in the TSNI study, this was only done for intervention II and not for intervention I. The last step, Maintenance, was also not measured as the TSNI study did run for long enough to measure this step. Hence, the current researcher has attempted to measure the following output steps (i.e. outcomes of this research): exposure, attention, comprehension, skills acquisition, recall, and attitude (change).

Measurement

Creating scales, indexes, or any measurement/assessment tool is part of the research process. Measurement is at the core of doing research as it consists of two basic processes: conceptualization and operationalization. [114, 115]

Conceptualization is the process of taking a construct or concept (i.e. the output steps) and refining it by giving it a conceptual or theoretical definition. This definition can derive from research questions or hypothesis and are then made into clear and consistent definitions which are then agreed-upon by others in the scientific community. Sometimes, the researcher comes up with a new conceptual definitions, however, this only happens when the researcher has an intimate familiarity with the topic. More common is the process by which a researcher notes agreements and disagreements over conceptualizations in the literature review, and ends up favouring someone's conceptual definition. Conceptualization is often guided by the theoretical framework that the researcher is committed to. Conceptualization is the same as asking: "What is meant by X in this research?" [114, 115]

Operationalization is the process of taking a conceptual definition and making it more precise by linking it to one or more specific indicators. These are usually things with numbers in them that reflect empirical or observable reality. So it specifies how a concept will be measured: "How was X measured?" [114, 115]

After these two basic processes of measurement, there is an advanced process which consists on the 4 levels of measurement: nominal, ordinal, interval and ration. The first 3 were the ones used in this study. [114, 115]

The nominal level of measurement describes variables that are categorical in nature. From all the types it is the one with less accuracy. In this research it was the level attributed to comprehension-lexical meaning. [114, 115] The ordinal level of measurement describes variables that can be ordered. Most opinion or and attitude scales are ordinal in nature. In this research it was the level attributed to attitude change. [114, 115] The interval level of measurement describes variables that have more or less equal intervals. In this research it was the level attributed to all the other output steps. This level is more accurate than the previous ones.[114, 115] Advanced statistics require at least the interval level. Variables should be conceptually and operationally defined with levels of measurement in mind since it is going to affect how well is that data analysed.[114, 115] After the 4 levels of measurement there is an even more advance phase, which consists of determining the validity and reliability of the measurements. [114, 115]

Validity is asking, "Are we measuring what we think we are measuring?" While reliability is the consistency of your measurement, or the degree to which an instrument measures the same way each time it is used under the same condition with the same subjects. It asks the question "How reproducible are the results?"[116, 117]

There are three main types of validity: content, criterion and construct

Content validity is related to sampling and asks the extent to which items in an instrument are reasonably representative of the larger domain or subject being measured. Face validity which is an aspect of content validity refers to the practice of having a panel of experts or potential users of the instrument review an instrument after it has been developed and making judgements about the extent to which a particular instrument measures what it is intended to measure. [52, 116]

Criterion validity compares data obtained to a criterion. There are two kinds of criterion validity: concurrent validity and predictive validity. In the case of concurrent validity, the criterion is occurring at the same time as the data is collected. For example, two methods that measure similar things, like the repeated 24-hr recall and food frequency questionnaires. In the case of predictive validity the criterion occurs in the future. For example, is it possible to use McGuire's sequence of steps to predict behaviour change? [52, 116]

Construct validity is about the extent to which respondents' scores on an instrument provide good measure of a specific construct. For example, the attention observation

method used in the current study was first administered to a sample of university students to give an idea of what could be considered good attention. [52, 116]

There are four general classes of reliability estimates, each of which estimates reliability in a different way:

Inter-Rater or Inter-Observer Reliability is used to assess the degree to which different raters/observers give consistent estimates of the same phenomenon. [52, 116, 117]

Test/retest is used to assess the consistency of a measure from one time to another. Reliability is high when one gets the same score on test 1 as on a repetition of the same test later on. A correlation between the two separate measurements is usually calculated. [52, 116]

Internal consistency estimates reliability by grouping questions in a questionnaire that measure the same concept. For example, you could write two sets of three questions that measure the same concept (say class participation) and after collecting the responses, run a correlation between those two groups of three questions to determine if your instrument is reliably measuring that concept. One common way of computing correlation values among the questions on your instruments is by using Cronbach's Alpha. Cronbach's alpha splits all the questions on your instrument every possible way and computes correlation values for them all. In the end, your computer output generates one number for Cronbach's alpha - and just like a correlation coefficient, the closer it is to one, the higher the reliability estimate of your instrument. [52, 116]

The primary difference between test/retest and internal consistency estimates of reliability is that test/retest involves two administrations of the measurement instrument, whereas the internal consistency method involves only one administration of that instrument. [52, 116]

4.1 Study population

For the current research study a sub sample of the TSNI population assigned to intervention I and II was randomly selected. All participants in the current research study were the biological mothers of the TSNI reference children.

4.1.1 Sample size calculations

Different sample sizes were used for the different outcomes/output steps. Hence, sample size calculations in the current research study varied among the output steps measured. Exposure was recorded for all TSNI research study mothers in intervention I and II, but for the majority of all other steps a sample size and sensitivity calculation was carried out in EXCEL using the following equation[118]:

$$n > \frac{\left\{ \mu \sqrt{\pi(1-\pi)} + \nu \sqrt{\pi_0(1-\pi_0)} \right\}^2}{(\pi - \pi_0)^2}$$

n = required minimum sample size

π = proportion of interest

π_0 = null hypothesis proportion

μ = one-sided % point of the normal distribution corresponding to 100% - the power

ν = percentage of the normal distribution corresponding to the required (two-sided) significance level, e.g. if significance level = 5%, $\nu = 1.96$

Sample size for attention and comprehension-lexical meaning output steps:

No studies were found that have measured attention in a similar way and study setting as the current research study. Hence, there were no data available on which to base the sample size calculation for this step. A pragmatic decision was therefore made to have a sample of sufficient size to detect at least a 20% difference in the prevalence of attention scores between the two approaches (i.e. 65% in individual class approach, and 45% in the group class approach), with a 5% significance level, 80% power, and two-tailed test. A sample size of 89 mothers was obtained for each approach (i.e. GC and IC). For the reason that understanding-lexical-meaning was measured at the same time as the attention measurements, the same sample size was used.

A major limitation is that when sample size was calculated it was done on the basis of achieving a difference between approaches (i.e. GC versus IC) and not between intervention groups, for that reason, GC intervention I sample size is less than half of what it should have been (i.e. $n=42$).

Post-hoc calculations were carried out to find out what would be % difference between the two intervention groups with the reduced sample size for GC intervention 1. With a

sample size of 42, which is the case for intervention 1, a difference of 30% between the two intervention groups would be considered significant.

Sample size for comprehension-psychological meaning, skills acquisition and recall output steps

The same task was used to measure these three output steps. Due to the nature of the skills acquisition task which was time and resource consuming, a pragmatic decision was made to have 40 mothers in each intervention group.

Unlike the sample distribution for the attention measurements whereby group class measurements were carried out for mothers in both interventions, measurements taken for these last 3 output steps in the group class approach were only taken for intervention I subjects.

Sample size for attitude output step

Since the TSNI nutrition education program contained many messages aiming at changing behaviours regarding breastfeeding and complementary feeding practices, the attitude tool was based on these two topics and for that reason it was thought that sample size should be based on other similar studies. No studies, however, were found that have compared attitudes to weaning and breastfeeding practices between different approaches/interventions. Hence, there were no data available on which to base the sample size calculation for this step. A pragmatic decision was therefore made to have a sample size of 127 mothers for each intervention. In addition, a control group (i.e. the TSNI control group) was also measured.

4.2 Exposure

Were the mothers exposed to the messages? According to McGuire's model, without this crucial step the ascent through the steps to behaviour change would be impossible, as people cannot be persuaded by a message they have not been exposed to.[77]

4.2.1 Example of other studies' methods for measuring exposure

Exposure is probably the simplest and the most straight forward of all the output steps. Measuring exposure would depend on the message characteristic, for example, if it was a visual display of some kind (billboard, leaflet/brochures, poster, TV and others), and on the mode of delivery used. It is measured slightly different in mass communication and in face-to-face communications.

In mass media/ communication campaigns it usually consists of an estimate of the numbers (or a proportion) of individuals who could have been exposed to messages or channels over a given period of time. There is no control over who was exposed and it does not necessarily require directly surveying the target audience. [82]

Exposure to face-to-face communications refers to the proportion of people, who are or were exposed to the message. It can be measured as the percentage or proportion of people that have attended or completed classes or counselling sessions. [119] [92] [93] This approach allows for more control over exposure than with mass media communications, leading to more precise results.

4.2.2 Methods for measuring exposure in this research study

Exposure to group classes was recorded by measuring attendance to several topics, for this reason it will be referred to as attendance rates. Overall there were 12 topics. Each of the 4 health promoters had a list of all TSNi research study mothers and before starting a class they recorded if the mothers were present or absent. Topics were repeated at least once and if mothers showed at least one attendance to a topic, then she was considered as *exposed* to that topic. If the mother missed both the first class for a topic and its repetition then she was considered as *not exposed* to that topic. Attendance to group classes was recorded for intervention I and intervention II mothers.

Exposure to individual classes was recorded by measuring attendance to the 6 sessions. Each of the 4 health promoters had a list of all TSNI research study mothers that were supposed to receive individual classes. Only when the health promoter successfully visited the mothers did she record that they had been exposed to the class.

4.3 Attention

Did the mothers pay attention to the message?

4.3.1 What is attention?

The most famous definition of attention was provided by one of the first major psychologists, William James: "Everyone knows what attention is. It is the taking possession by the mind in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought...It implies withdrawal from some things in order to deal effectively with others." [120]

4.3.2 Attention and learning/performance

According to Zeaman and House's hypothesis the learner must first attend to the task at hand or relevant stimulus, before responding to it. Until appropriate attention occurs, learning does not take place. [121, 122] Hence, attention is an integral part of learning and inadequate attention results in learning deficiencies. As suggested by Piontkowski (1979) [123]: "It is hard to believe that the student who disregards instruction will benefit from it. Thus, the teacher needs reliable signs of the student's general state of attention." [123]

The most important types of attention required for learning in a classroom setting are selective and sustained attention. [124] Selective attention involves focusing on a specific aspect of a scene while ignoring all other aspects. In a classroom situation, selective attention enables the student to overcome certain disturbances, such as, noise coming from other students, and concentrate on the signal of interest (i.e. what the teacher is explaining). [122] Sustained attention (also referred to as vigilance) is "the ability to direct and focus cognitive activity on specific stimuli". [125] In order to complete any cognitively planned activity one must use sustained attention. An example is the act of reading a newspaper article. One must be able to focus on the activity of reading long enough to complete the task. Problems occur when a distraction

arises. A distraction can interrupt and consequently interfere in sustained attention. An individual's focus of attention can only last 30 seconds, it digresses but can be re-focused at will.[125]

The act of learning intimately involves the auditory and visual senses. "While both senses are important, it is reasonable to suggest that under normal circumstances the eyes are required to input greater amount of information during learning". [122, 123] Because eye movements reflects a person's allocation of attention in a scene, it can be used as an indicator in the attentional system.[122, 123]

4.3.3 Eye movements

There are two main eye movements relevant to attention: Saccades which is the principal method for moving the eyes to a different part of the visual scene, and are sudden, rapid movements of the eyes, and, fixation or gaze which refers to maintaining the gaze in a constant direction. The direction of gaze carries information about the focus of the user's attention. Humans (and other animals with a fovea) constantly alternate saccades and visual fixations.[126, 127]

4.3.4 Methods for measuring attention

Researchers have developed sophisticated laboratory techniques/methods to measure reactions to attention. The development of sensitive and reliable instruments for the measurement of eye movements dates from the 1960's. However, almost all are lab-based, controlled and invasive and therefore would not be appropriate under natural "real world" conditions.[123][122] Nonetheless, the 1960's-1970's methods of observation, as reported in the literature, also included the subjective estimation of eye position by an independent observer, a simple method, non obtrusive, even though less precise.[122] In a classroom setting under natural conditions a similar method would be preferred – to measure eye movement activity by monitoring eye fixation.

In school educational research the only thing that is of interest is to monitor, in a general way, if the student is attending to what the teacher is saying/showing, more precisely, if he/she has sustained attention. The interest is to monitor if the student is gazing, fixating or making eye contact with the teacher. Therefore we do not need very sophisticated and expensive, and for sure, not invasive techniques. "Fortunately, two simple methods, practical for the classroom are also valid: a) notice where the

student's eyes are directed, and, b) ask the student immediately after the event what is on his mind.”[123]

Like all measurements, measuring eye-fixation in a classroom as a measure of attention has many drawbacks:

1. Accurate measurement proves to be very difficult in a natural setting.
2. Shifts of attention can be made overtly by moving the eyes or covertly with attention being allocated to a region of space that does not correspond to the current direction of gaze.[128] This is to say that, what we look at may not be what we attend to. It is possible to look in one direction but actually notice changes in another direction. Salvucci (2000) argues that “although eye movements and visual attention are certainly correlated, they do not always correspond directly, especially for complex visual stimuli. This separation of visual attention and eye movements can create great difficulties when attempting to evaluate cognitive models based on their predictions of visual attention.”[127] However, when measuring general attention in a classroom setting we are not trying to measure complex visual stimuli, nor evaluate cognitive models, and also, the examples that Salvucci gives for when this problem would be relevant is when two objects are near each other, or when reading, words or phrases are near each other and therefore, the experimenter cannot monitor for sure the location where the subject is fixating. Nonetheless, some studies have suggested that, though humans and primates can look in one direction but attend in another, there is underlying neural circuitry that links shifts in covert attention to plans to shift gaze. So, if we attend to the right hand corner of our eye, we ‘want’ to move our eyes in that direction, and have to actively suppress the eye movement that is linked to this shift in attention. [128-130]

4.3.5 Example of other studies’ methods for measuring attention to health messages.

Not many studies have attempted to measure attention to health messages. “In persuasion research, researchers generally believe that attention is a necessary (but not sufficient) condition for effective communication. However, persuasion researchers have largely ignored attention as a dependent variable in evaluation studies. In contrast, attention has been extensively studied in cognitive science and in neuroscience for different modalities, such as visual attention, auditory attention, or cross-modal processing.”[90]

McGuire, in 1968, when thinking on how to measure attention wrote the following: “It is usually hard in practice to distinguish operationally between the steps of attention and

comprehension. We can, however, measure them jointly by administering a reading comprehension test that measures how well the person can report the message content, as compared with a control subject who has not been exposed to the message. His score on this test of recall provides a measure of how well he attended to and also comprehended the contents, without clarifying whether failures of recall were due to inadequate attention, inadequate comprehension, or both. Perhaps new techniques (such as pupil dilation) will someday allow us to get a direct measure of attention and thus allow us to tease out the relative contribution of these two mediators.”[70] In 1985, he wrote: “...measured by contemporaneous eye movements or EEG potentials, by subsequent ad recognition tests, etc.”[71]

Of the studies that have attempted to measure attention to health messages, they have done so by :

1. Recording event-related potentials and reaction times which are generated from a continuous measure of brain activity by means of an electroencephalogram. [90]
2. Using eye-tracking to study how participants view health messages. It consists of wearing a light weight optical component mounted on a headband with a fitted video camera. The pupil and corneal reflection information was videotaped and the signal containing the field-of-view was simultaneously recorded. The outcome measured was *eye fixations*. [131]
3. Asking subjective questions, scoring them and then summing the scores. For example, asking subjects if they remember receiving and then reading the magazine (i.e. the health communication material), then scoring the response (where 0 = no recall, 1= prompted recall, and 2 = unprompted recall and 0 = reading none of the magazine, 1 = reading some of it, 2 = reading most of it, and 3 = reading all of it; or using a Likert Scale for the same purpose), and then combining scores resulting in a composite attention variable. [87] [88] [89]
4. Asking a subjective question. For example: “Have you seen any signs in bars, restaurants, stores, or other places that warn people about drinking?”[132] [76, 133]
5. Asking more specific questions, like for example, if referring to a poster, what the background colour of the poster was, whether there was any wording and if so, what it was, whether there was a picture and if so, of what. [134]

6. Asking individuals who had looked at a series of messages whether they recall having seen or heard message X. Bogus messages could be placed on the list to try detect falsely affirmative responses.[82]

7. Systematically observing subjects, moving from one subject to another at 5s intervals. As each subject is observed, the experimenter notes whether or not the subject is attending to the message (ex. a videotape). [135]

8. The Paediatric Examination of Educational Readiness (PEER), which has been compiled to enable health providers to make direct developmental and behavioural observations of children. One component of the PEER is an attention activity rating scale which allows direct observations of attention-activity at specified times during the testing. During these observation several attentional signs are observed and rated, including distractibility (the extent to which a child is "called away" by auditory, visual, or inner-personal stimuli), and attentional absences (the number of times a child's mind clearly wandered from the situation at hand).[136]

As can be seen from the above, attention to health messages has been assessed using a mixture of methods. The first two methods would not be applicable and appropriate under natural conditions as they are quite invasive and also, it is difficult to know whether mechanisms that appear stable in a controlled environment will continue to exist outside that environment. The third and 4th methods seem to be confusing attention with exposure. Moreover, because questions are subjective it is difficult to ascertain if subjects did indeed pay attention. The last four methods seem to be the most appropriate for a natural setting. Only 2 studies[135] [136] measured attention using direct systematic observation in natural settings. And as discussed by Kingston: "These observations will form a much needed personal-level description of cognition as it operates in real-world settings." [137]

4.3.6 Methods used for measuring attention in this research

Attention measurements were carried out at a series of different group classes and individual classes. Mothers were selected in the following way: the researcher had a rota for the supervision of the 4 TSNI health promoters every week, and during these times, the attention measurements were taken. The researcher was not aware of which women's groups were going to receive the nutrition group class, or which subjects were going to receive the counselling session. Before the group class, subjects were selected at random according to where they were seated, not necessarily from the front

row, but scattered. Usually in group classes, 5-6 subjects were observed per class, and these would have to be seated in a way that it was easy for the researcher to see their faces.

In this study attention received two operational definitions (as detailed below):

Method 1 was based on the fact that more elaborated studies (i.e. lab based studies) measured sustained attention looking at eye fixation/gaze. This was adapted to the settings of the current study whereby sustained attention was defined as maintenance of eye contact between the mother and the health promoter or with the visual aid for 5 seconds at systematic observations. The method used here was very similar to the one described by Peterson *et al* [135]. Method 2 was based on the theory of selective auditory attention, where, subjects selectively pay attention to one auditory message instead of another. The method used here was similar to the one proposed in the personal communication from Elder. In this study, subjects were asked to decide if specific words were either mentioned or not during the class. The following is a detailed description of both methods used.

Method 1 – sustained attention observations (tool can be seen in appendix 4.1 and 4.2): This was based on direct systematic observations, where by the researcher observed eye contact between the mother and health promoter/ visual teaching material for approximately 5 seconds. This was carried out during both group and individual classes, which normally lasted between 30-45min. Observations were taken very discretely so as not to raise any suspicions from the participants. Overall, 10 observations were taken per mother and these were carried out at specific intervals through out the class. Observations were taken more or less every 2-3 minutes for the group classes and in the individual classes they were taken approximately 1 minute after the health promoter had started explaining each key message. In group classes the 5-6 mothers were observed systematically, starting with the first mother and moving to the next mother at 5 second intervals. As each mother was observed, the researcher noted whether or not the mother was looking at the health promoter or the visual aid. While this was simpler for the group classes, the pilot study showed that in the individual class setting mothers behaved differently than in a group class setting. The pilot study showed that in an one-to-one setting the mothers feel more timid which leads them to sometimes not make eye contact with the health promoters. Thus, the researcher had to consider other indicators of attention. This was usually nodding with the head as a sign of agreement to what the health promoter was explaining, the laughing and the agreement by saying “yes yes”.

Method 2- attention recall questions (tool can be seen in appendix 4.1 and 4.2): Very similarly to the Elder[82] and Cole *et al* (which asked participants if they had seen any wording in the poster display) [134] studies, in the current research study, mothers were asked, immediately after the class, if they recalled having heard word X (during the class). Two recall questions were asked, one being a bogus question in order to detect falsely affirmative responses. Questions were of the kind “during this teaching session/class did you remember hearing the health promoter mention -----?”

All measurements were taken by the researcher, but recall questions were asked in the local language and this was carried out by the health promoters who had received extensive training on how to ask the questions. Nonetheless, this process was supervised by the researcher at all times who would also write down the answers.

Considering that nutrition education in a rural setting of a developing country takes place under less than ideal conditions, it is important to examine how commonly occurring factors might affect learning. For example, during the attention measurements in individual classes, 73.6% of subjects were holding their child in the class. This was also seen in the group classes even though it was not recorded. The child would often be seeking the mother’s attention. Child disturbance in group classes was measured in a different way to that in individual classes. While in the individual classes child disturbance was measured on each observation, in the group classes the observer only observed the total number of times that any child cried during the class. Because whenever a child cried or needed the mothers attention during a group class the mother would stand up to attend to the child, the number of times that someone stood up was measured as this could also have been a potential distracter. For these and other reasons, factors that might affect attention (and hence, comprehension) in this study were measured. These are: child’s disturbance during the individual classes, total child disturbance in group classes, other major disturbances in group classes (ex: women standing up in the middle of the class), duration of individual and group class sessions and class size. All of these present forms of distraction for the learner.

Scoring

Scores were allocated for each attention observation. A score of 1 was allocated for each observation where the mother was found to be making eye contact with the health promoter/visual aid and a score of 0 if she was not. The maximum possible score for each mother was 10.

Attention recall questions were also scored. A score of 1 would be awarded to both an affirmative response to the non bogus question and a negative response to the bogus question, hence, mothers could get a maximum total score of 2.

The assessment of each method was kept separate.

Piloting

Pilot testing was carried out to test and further develop if necessary the attention tool (i.e. the two attention methods). This was carried out on the TSNI study areas on 10 IC mothers and 4 group classes. Ambiguity and clarity of attention recall questions were tested. Processes within both approaches were observed. Timing of attention observations were tested. Number of attention observations were tested. Initially, 5 observations were being made and because this number was found to be inappropriate it was increased to 10.

At the same time, another pilot test was carried out, in a completely different setting. The aim was find out a cut off to what could be considered good attention and bad attention. Measurements were taken in two classes in a polytechnic school in the capital of Zambezia Province. Overall, 10 students were observed. The average attention level was 6 and hence, this was considered the cut off point. The reason why this setting was assumed to be a good reflection of good attention standard is because students in a university level have been more exposed to education, classes are more comfortable, there is less disturbance, and teaching is better.

4.4 Comprehension/Understanding

Did mothers understand/comprehend the messages?

Once an individual has oriented himself/herself towards the message, he or she will begin to recognize that pictures, objects, or words are present in the message and will access the stored meanings of those words.

In education and psychology, comprehension is synonymous to understanding, therefore both terminologies will be used here interchangeably. [138]

Why is comprehension important in education? As mentioned in the last section, paying attention to a class or message is important for education to be effective and until appropriate attention occurs, learning does not take place. Understanding the

class or message is also as important as paying and maintaining attention. So, for one to understand a communication it is a requirement that that person has first paid attention to it. However, paying attention to a class or to someone reading a text does not necessarily imply that one will understand the information. Take for example this text that Bransford read to his subjects:[139]

“The procedure is actually quite simple. First you arrange items into different groups. Of course one pile may be sufficient depending of how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step; otherwise, you are pretty well set. It is important not to overdo things. That is, it is better to do too few things at once than too many. In the short run this may not seem important but complications can easily arise. A mistake can be expensive as well. At first, the whole procedure will seem complicated. Soon, however, it will become just another facet of life. It is difficult to foresee any end to the necessity for this task in the immediate future, but then, one never can tell.”[139] Even if subjects paid and maintained attention, they would find it very difficult to understand it, although it was possible to understand the individual sentences in the passage. It was the relationship among the sentences that was unclear, rendering the integration of the idea units impossible.

4.4.1 The Nature of Understanding

The study of comprehension is not as documented and researched as for example that of attitudes, attention, memory and recall, and even skills acquisition. This could come from the fact that it is such a multi-dimensional, complex and blurred concept, with no distinct line dividing its beginning and end, making it many times difficult to distinguish and separate from other constructs.[140]

Defining comprehension is not an easy task since it is difficult to find a definition that would encompass all the facets of so complex a concept. Yingxu and Gafurov (2003) [141] define it as: “Comprehension is a higher cognitive process of the brain that searches relations between a given object or attribute and other objects, attributes, and relations in the long-term memory, and establishes a representational model for the object or attribute by connecting it to appropriate clusters of memory. It is recognized that although knowledge and information are powerful, before any information can be possessed and processed, it should be comprehended properly.”[141]

Long (1989)[142] world knowledge, and filling in the gaps with logical guesses. Intake is then recoded and stored in long-term memory in the form of propositions, or basic

meaning forms.” [142] describes the process of understanding as: “Learners construct meaning during the comprehension process by segmenting and chunking input (i.e., that which they hear or read) into meaningful units. Actively matching the results, known as intake, with their linguistic and world knowledge, and filling in the gaps with logical guesses. Intake is then recoded and stored in long-term memory in the form of propositions, or basic meaning forms.” [142]

White and Gunstone (1992)[140] break down the process of understanding a concept into several types of knowledge: *propositions*, *strings*, *images*, *episodes*, and *intellectual skills*. In addition to these six types of knowledge, another more general one can also be added- *cognitive strategies*. These are broad skills used in thinking and learning, such as being able to maintain attention to the task in hand, deducing and inducing, perceiving ambiguities, sorting out relevant from irrelevant information, determining a purpose, and so on. [140]

As can be seen from the above definitions, knowledge plays a vital role in understanding and also when a reader comprehends a text, its content becomes part of the reader’s overall knowledge.[143] For effective understanding to occur there must be a relationship between inputs and current activated knowledge of the learner. So, everyday comprehension depends on people’s abilities to make inferences and assumptions based on their general knowledge.[139] Though the concept of understanding is widely employed in educational discourse, and in most of the context, understanding is many times ambiguously regarded as a kind of knowledge, or more specifically, theoretical knowledge.[144] It is however believed that knowledge is the simple awareness of bits of information while understanding is the awareness of the connectedness of this information that allows knowledge to be put to use and therefore represents a higher level.[138]

There also seems to be a very blurry division between understanding and skills acquisition, as if, acquiring skills is a sign of comprehension, even though McGuire distinctly separates the two in his model. White and Gunstone, include skills acquisition, calling it intellectual skills, as a component of understanding. They write: “they are memories of procedures which gives one the capacity to carry out classes of tasks, such as being able to substitute values in a formula, or apply a formula to solve a problem. It is important to distinguish between the verbal statement of the procedure and the capacity to perform it. For example, if someone asks how to distinguish cats from dogs, one could answer with a verbal knowledge, propositions, which one could

possess without actually being able to use it to tell cats from dogs. Being able to distinguish the two is the intellectual skill. [140]

Comprehension does not seem to have a unidirectional linear relationship with memorisation and recall/ retrieval like in McGuire's model. In the afore-mentioned study by Bransford, after reading the incomprehensible text, subjects were asked to recall it. As a consequence of not understanding it, recall scores were quite low.[139] It might therefore be tempting to assume that there is a unidirectional linear relationship between comprehension and memory/recall, and that the better the comprehension, the better the memory performance will be. Although putting information into context that facilitates understanding may sometimes help memory (i.e. it is easier to understand Bransford's text if we know that it is referring to washing clothes), the use of prior knowledge to make information coherent and sensible can sometimes result in distortions in recall.[139] Research reveals that the ability to memorize facts does not necessarily imply understanding of a concept. Apparently, understanding is not so much a destination as it is a point along a continuum. We may never arrive at perfect understanding of a subject, but we can deepen our understanding beyond the superficial. At some point, learners' grasp of the concept becomes deep or sophisticated enough that they can use their knowledge in a practical way. It is this level of comprehension that teachers aim for. That is, *understanding* can be defined as "the ability to apply a new concept outside a textbook or classroom context, use a concept to make sense of complex, real-world situations, or express a concept in a meaningful way to others." [145] Reeder, a maths teacher gives an account of her experience with her geometry students. She had just taught her students the formula for calculating the length of an arc of a circle and found that her students could define it and use it in the exercises from the textbook. She then asked her students to do a real life exercise, which they would have to use the just-learned formula. But she did not reveal this to them. What she found is that none of them employed the formula they had just learned.[145]

4.4.2 Example of other studies' methods for measuring comprehension to health messages.

Measuring comprehension is not an easy task either. The tests of fifty years ago are largely measures of ability to recall facts or to apply standard algorithms, while those of today more commonly make an effort to see whether the knowledge can be used to solve novel problems.

A review of studies that have measured comprehension to health messages retrieved many studies that have measured reading comprehension and have used state-of-art techniques, such as the modified Cloze procedure [146]. However, the current research was carried out in subjects with hardly any literacy levels and therefore assessing reading comprehension would not be appropriate.

More appropriate methods would be:

1. To measure by using a 7-point Likert scale (very easy to not at all easy) to measure the degree of ease to understand the messages.[87]
2. To ask questions regarding the content of the education. [147]

The problem with the first method is that it is subjective and so it is difficult to ascertain if subjects did indeed understand the health communication. Many studies confuse comprehension with knowledge.[86] [82]

The best method would be the second one, but once again many studies tend to measure recall instead of comprehension. Houts *et al* [76] explain the possible misuses of the construct comprehension: "Comprehension is the process of interpreting the meaning of words or pictures to understand their collective meaning. It is different from recall which is the process of retrieving individual words or picture elements from memory. People may remember material without, necessarily, understanding it. Therefore, for a study to qualify as evaluating comprehension, the outcome measure had to go beyond simple recall and ask respondents to explain or do something with the information presented."[76]

Rimer and Glassman (1984) [78] study go into more detail on certain aspects of comprehension. They argue that in order to measure comprehension to health messages, it should be defined in terms of those elements that can be tested. They sub divide comprehension into 3 categories: [78]

Psychological meaning—refers to the learner's ability to relate the information to his/her life and to understand its personal implications, for example, to indicate what the personal effect of regular exercising would be. In their study the psychological meaning question was: "What would be the benefits and drawbacks for you of starting an exercise program?"[78]

Operational meaning—refers to the person's ability to use the information, for example, to design a personal aerobic fitness program or calculate a target heart rate. In their

study the operational meaning question was: "What is your target heart rate or how would you find it?" [78]

Lexical meaning- refers to definition of the key concepts and words of a health education message, such as target heart rate. This is the level of comprehension assessed most frequently in health education programs. In their study the lexical meaning question was: "Tell me in your own words what does these words mean: (1) aerobic, (2) target heart rate." [78]

McGuire, when discussing tests to measure the comprehension step, mentioned that it could be measured by using *posteriori* recall or recognition tests of message contents which would be the equivalent of comprehension-lexical meaning.[71]

4.4.3 Methods used to measure comprehension in the current research

In the current research, comprehension was measured following the same principle as Rimer and Glassman's (1984) study where comprehension is assessed by dividing it into its lexical meaning and psychological and operational meaning.[78] Lexical meaning was assessed separately from psychological and operational meaning and the last two were combined into one.

4.4.3.1 Comprehension lexical meaning

Comprehension lexical meaning was assessed right after the attention measurement, at the end of the sessions and group classes. Hence, the same subjects measured for attention were the same measured for comprehension- lexical meaning (*tool can be seen in appendix 4.1*).

Comprehension psychological-operational meaning was measured in another occasion, at the same time as recall and skills acquisition and hence the same subjects were measured. Overall there were three lexical-meaning questions, but only one was given to participants at each session/class. These questions were designed according to the content of the group and counselling classes and had to be related to that day's class topic.

In order to avoid different results for both approaches due to different questions, these questions had to be the same. Hence, after reviewing possible questions and testing

them, the following three lexical-meaning questions were developed for both group classes and individual classes:

- 1) *What is the source of the pathogens that cause diarrhoea?*
- 2) *Overall, how many main meals a day should this child (i.e. showing a cultural accepted picture of a child of approximately one year old) have?*
- 3) *Give examples of 3 vitamin A-rich foods.*

The correct answer for the first question should have been: *faeces*. [148] In group classes this question was asked after the hygiene and diarrhoea topic was presented, where participants were explained the f-diagram [148]. In the individual classes the question was asked at the end of the classes, and it was related to one of the key messages mentioned in every classes.

The correct answer for the second question should have been: *3 main meals or 3 main meals: breakfast, lunch and dinner*. Again, in group classes the question was related to the topic of the class, which was about weaning, and it was one of the key messages given in every individual classes.

The correct answer for the third question could have been any three foods from the following list: orange fleshed sweet potato, eggs, pumpkin, small fish, dark green leaves, mango, papaya and carrots. In group classes, this question was asked in those classes where this list was discussed and in the individual classes it was related to one of the key messages mentioned in every classes.

The above three questions are considered lexical meaning type of questions because they are more of the recognition and short-answer recall type questions.

Scoring

The tool was assessed using a scoring system whereby 0 = incorrect answer and 1= correct answer.

Piloting

Piloting was carried out at the same time as the attention piloting. The pilot tested the clarity of the question and checked the answers given. Questions that lead to totally unexpected answers were reviewed for their clarity. Lexical meaning questions were administered in the local language by the health promoters. The language used after the translation was much simpler than the version mentioned above, so for example,

the word *pathogens*, which does not exist in the local language, was translated into something more like bugs or with a similar meaning. The health promoters themselves translated the questions and these were standardized for all of them. Even though it was the health promoter that would ask the question, the researcher was supervising this process at all times, and would write down the answers.

4.4.3.2 Comprehension psychological-operational meaning

A task was developed to measure comprehension psychological-operational meaning, recall and skills acquisition.

This task consisted of intervention I and II mothers first participating in cooking demonstrations where they were taught how to prepare weaning food recipes using orange-fleshed sweet potato and other nutritious and locally available foods and then followed up 1.5 to 2 months afterwards where they were asked to chose one of the recipes and prepare it on their own. The following is a description of the composition of the 3 recipes taught in each locality:

Locality	Recipe 1 ingredients		Recipe 2 ingredients			Recipe 3 ingredients			
Malei	OFSP	Coconut milk	OFSP	Toasted peanuts		OFSP	beans	Coconut milk	Dark green leaves
Posto Campo/luala	OFSP	Coconut milk	OFSP	Toasted peanuts		OFSP	fish	Coconut milk	Dark green leaves
Catale	OFSP	Coconut milk	OFSP	Toasted peanuts		OFSP	fish	Coconut milk	Dark green leaves
Mexixine	OFSP	Coconut milk	OFSP	Coconut milk	egg	OFSP	fish	Coconut milk	Dark green leaves

Intervention I mothers were recruited using the following procedure: The health promoters selected TSNi women groups from those available in each of the 4 localities. In each of these women's groups, every TSNi research study mother that attended the demonstration was recruited to be part of this task. Overall, 40 mothers needed to be recruited from the women groups, hence if the mothers that attended the demonstration were not enough to complete the required sample size of 40, then more women groups were recruited until more than 40 mothers (i.e. in case some decided to drop out the extra recruited could take their positions) had been recruited. Overall, 8

women groups participated in the demonstration. The sampling procedure for intervention II subjects was much simpler. Forty five mothers (i.e. extra subjects were recruited in case of drop outs) were selected at random from two villages in Malei and other two villages in Posto Campo/Luala and participated, individually, in the same cooking demonstrations as intervention I mothers.

Measurements for comprehension psychological-operational meaning were taken during the follow-up visit, specifically after the mothers were asked to prepare one of the recipes. (*Tool can be seen in appendix 4.3*)

The comprehension psychological-operational meaning questions consisted of the following:

- 1) *For each of the ingredients used in the chosen recipe, the mothers were asked to recall to which food group did it belong to and then were asked: **Why is this food group good for the child?***
- 2) *If the ingredient belonged to the protein or fat- rich food groups then they were asked: **what other ingredient from the same food group could you use if you could not get this one?***
- 3) *They were asked to recall how many recipes were carried out in the initial demonstration and then were asked: **Which one of the recipes is best for the child?***
- 4) **Why is it the best for the child?**

While questions 3 and 4 were only asked once, the number of times questions 1 and 2 were asked depended on the recipe chosen by the mother. Hence, a 4-ingredient recipe would receive 4 times question 1 and 2 times question 2, a 3-ingredient recipe would receive 3 times question 1 and 2 times question 2, while a 2-ingredient recipe would receive 2 times question 1 and only once question 2.

Unlike the understanding-lexical meaning questions which were asked straight after the classes and were related to the topic of the class (and therefore one cannot say if the mothers really understood the content or just memorised and retrieved what was said), some of the information required to answer the understanding-psychological-operational meaning question had been communicated 1.5-2 months before. Hence, these represented instead an aggregation and application of their understanding since the beginning of the program. They refer to the learner's ability to relate the information to her life, to use the information and to understand its personal implications. As said by Rimer and Glassman: "much of health education aims to teach skills rather than abstractions. The cognitive acquisition of skills, perhaps even more than the cognitive

acquisition of abstractions, requires that the information be linked effectively with the subject's prior behavioural experiences and the real-world conditions under which the skill must be practiced".[78]

Scoring

The assessment of the comprehension psychological-operational meaning was not as straight forward as that for the lexical meaning. They were also based on scoring the answers with either 0 = incorrect answer and 1= correct answer, which were then summed up. However, because different recipes entailed different number of times that questions were asked, the total score had to be a proportion of the total number of questions asked. So for example, if one mother had chosen the recipe with 4 ingredients, which means 8 questions overall, but she only managed to correctly answer 4 out of the 8 questions, she would then receive the score 4/8. This was then converted to percentages. Even if a mother got the recipe wrong by adding or omitting ingredient(s) she would still be questioned on the recipe she chose, regardless of being correct or not.

Comprehension psychological-operational questions were also administered in the local language by the health promoters. The health promoters themselves translated the questions and these were standardized for all of them.

4.5 Skills acquisition

Did the mothers acquire the new food preparation skills?

"Cognitive skill acquisition is acquiring the ability to solve problems in intellectual tasks, where success is determined more by subjects' knowledge than by their physical prowess."[149] Even though this description is looking at skills acquisition in a more cognitive way it can be linked to practice: "...complex actions, such as putting on a condom correctly, negotiating with a partner that a condom be worn, or being a peer leader require complex verbal and motor skills".[82]

For this step McGuire (1985)[71] wrote: "skills acquisition is involved when persuasive impact requires not only learning what but also learning how, as when a public health campaign against smoking must not only motivate but also convey skills for complying." [71] When discussing tests to measure this step, he mentioned that it could be

measured by using posteriori recall or recognition tests of message contents. [71] However, this seems to be more like comprehension-lexical meaning.

In a developing country setting live demonstrations should probably lead to better learning and skills acquisition than only giving oral information. This is supported by Nagarajan *et al* [150] who regarding oral rehydration solution (ORS) preparation discussed that: “rural women may not understand printed instructions properly, may misinterpret pictorial instructions and may not comprehend or forget easily verbal instructions”. In their study they compared a group that received only verbal instruction on the preparation of a ORS with another group that received the same verbal instruction and a live demonstration. Results showed that the group that received the demonstration could prepare a ORS more similar to the ideal concentration of ORS than the group that received no demonstration.

Of the studies mentioned in the literature review about complementary-feeding promotion studies in developing countries (see chapter 1) only 3[26, 47, 49] carried out demonstrations on the preparation of weaning foods. However, none of them measured skills acquisition. None measured if mothers after being taught were actually able to execute the porridges on their own.

It seems important to measure skills acquisition in this way because it shows the result of the subject's learning process. The aim of these practical sessions was to teach the recipes and how to prepare them. What good would it do, if the mothers stored the information but then could not apply, in a practical way, this knowledge?

4.5.1 Example of other study's methods for measuring skills acquisition in health communications.

The following is a review of methods used to measure skills acquisitions within nutrition education. It also includes studies that have used cooking demonstrations since it is more relevant to the current research.

Elder suggests that one should use direct observation to measure skills acquisition.[82] In the study by Bull *et al* (2001), subjects were asked to view printed health education materials (HEMs) on weight loss. One-month afterwards, they measured skills acquisition during a telephone interview by measuring the numbers of items participants stated they had learned from the HEM. [87]

The study by Taylor *et al* [151] (2000) measured skills acquisition in a nutrition education program for healthy lifestyles in a low-income Hispanic population. Skills taught during the program included the discussion of basic nutrition knowledge to tips on choosing healthy foods. A questionnaire was used to measure skill acquisition, and then scored using the percent of total questions answered correctly. [151]

A study by Anderson *et al* [152] (2002) developed a questionnaire designed to assess knowledge, including skills, of applied nutrition in children participating in an after-school care dietary intervention programme. Among others, the questionnaire items assessed knowledge of basic food preparation and cooking skills for preparing familiar dishes. The part of the questionnaire which assessed cooking skills consisted of open-ended questions. Questions would be of the kind: "what are the main ingredients needed to make coleslaw?" Each correct answer was awarded one point and the possible maximum score was 14. [152]

A study by Hotz [49] used demonstrations to teach mothers in rural Malawi how to prepare thick porridges (i.e. more energy dense), how to reduce phytate content of maize and how to enrich plain maize porridges with other ingredients. This was evaluated using questionnaires whereby the mothers were asked about the preparation methods. This was referred to as Knowledge in the study and not skills acquisition. Positive scores were assigned to correct knowledge of preparation methods. [49]

Another study by Giere *et al* [153] (1989), aimed at teaching meal planning, meal preparation and shopping skills to individuals with moderate and severe handicapping conditions. In order to measure skills acquisition, pre-and post-tests were used. This measured preparation skills: participant's ability to turn the stove and oven on, set a kitchen timer to different times and to use measuring cups. Also measured meal preparation and planning. Following the pre-test, individuals received training until mastery. Following mastery, the subjects were then asked to prepare each of the meal items. The trainer would then record the number of steps performed and whether the individual could correctly accomplish the task. Tasks were hence evaluated by looking at the percentage of steps completed independently. [153]

The above studies included a variation of questionnaires and participant observation for measuring skills acquisition, but, agreeing with Elder, the best method would be the one used in Giere's study and which shall be used in this research study. Participant observation is a preferable method to assess skills acquisition than questionnaires.

Hotz himself mentions that the thickness of porridge could not be measured using a questionnaire, only by observing its viscosity.[49]

Only 1 study [96] was found that compared skills acquisition between group and one-to-one approaches. In this study mothers were taught how to prepare ORT in a group of 5 or individually. A year later they were asked to prepare the ORT on their own. Sodium concentrations were compared between the two groups and results show that there was no significant differences between the two groups. [96]

4.5.2 Methods used to measure skills acquisition in the current research

As discussed above, it is better to use demonstrations than just give oral information. Especially, if the program is teaching mothers about weaning foods. But even more important is to assess that the mothers have indeed gained any skills in the preparation of weaning foods. This was the case in the current research.

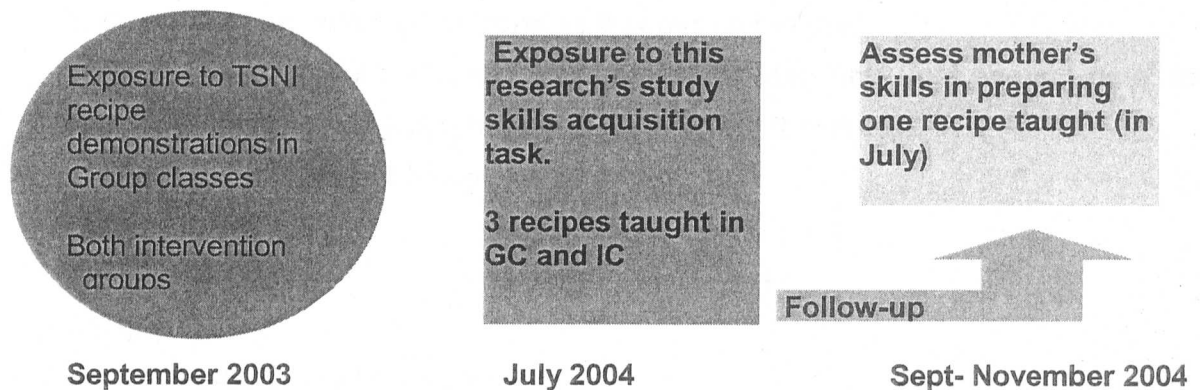
The reason why the TSNI found it important to include cooking skills demonstrations in the nutrition education program was in part to show mothers how to integrate OFSP in the children diet but also because baseline data and other previous studies have shown that weaning foods were based mainly on a porridge made of cassava flour and salt or cassava flour and sugar. This consisted in cooking the cassava with water and then adding the sugar/salt. Before introducing the OFSP, the population was already eating the white-fleshed variety. The method of cooking was by simply boiling the potato, chopped into pieces if it was a big piece. No attention was paid to the cooking time or the peeling.

The description of the methods used to measure comprehension psychological-operation meaning give an idea of how this output step was measured, but it shall be explained in more detail. (*Tool can be seen in appendix 4.3*)

In both group and individual recipe demonstrations, mothers were taught how to prepare 3 recipes step-by-step (i.e. porridge for children using OFSP and other nutritious and locally available ingredients) and these ranged in cooking method (i.e. normal boiling and steam cooking), difficulty and number of ingredients. Specific guidelines for the preparation of each recipe were given during the demonstrations. Following the sampling procedure already mentioned, these same participants were

then followed-up approximately 2-4 months later (see diagram below). On this occasion, the mother was contacted and reminded briefly of the initial demonstration. She was then invited to choose and prepare one of the recipes. Neither the health promoter or the observer disclosed any information regarding which recipes were carried out at the initial demonstration and their ingredients. The mother had to recall on her own how many recipes were carried out, choose one to carry out a week later and recall the ingredients used in her chosen recipe in the initial recipe demonstration. The health promoter then confirmed again the recipe chosen by the mother with the ingredients that the mother recalled, regardless of being correct or not. The study would only provide those ingredients chosen by the participant that she could not get herself. A date was arranged when the skills acquisition task would be carried out.

Time line:



This same procedure was carried out for all 80 mothers, 40 from Intervention I and another 40 from intervention II.

All 80 mothers were to be assessed individually, even those 40 intervention I mothers that had received the initial demonstration in a group setting. However, as can be seen from the list of topics taught at group classes and the diagram, above, both intervention I and II mothers had already been exposed to these same recipe demonstrations before.

On the day of the assessment, the mother was given the ingredients that she had chosen and could not find herself and asked to prepare the recipe. Both the health promoter and the researcher did not help the mothers in anyway, not even when the mother seemed a bit lost and needed help to remember a few steps. On these

occasions, the mother was told not to worry and to try and carry out what she remembered of the demonstration.

This is the second output step that has employed direct systematic observations (apart from the attention step). The researcher had a checklist of those steps (i.e. 6 overall), the specific guidelines, for the preparation of that recipe, that were explained (and demonstrated on the initial demonstration, and this was used to assess the skills acquisition task. The following is a list of the steps that the mothers should have followed. The steps differed for the cooking method used:

Steam Cooking:

- 1. Cooking method correct?**
- 2. Add only a bit of water;**
- 3. OFSP should be thin sliced**
- 4. Cover the pan with a lid as soon as it is put on the fire;**
- 5. While cooking, do not uncover the pan for a long time, just enough time; to check if it needs more water or needs stirring or if it is ready;**
- 6. The OFSP should not be overcooked**

Boiling:

- 1. Cooking method correct?**
- 2. The water should not cover the OFSP;**
- 3. Should boil the water before adding the OFSP;**
- 4. Should choose OFSP of the same size, or cut the bigger ones to have the same size as the small ones, or but the smaller ones in top of the bigger ones; However, it is better to have bigger pieces than smaller ones.**
- 5. Should not peel the OFSP**
- 6. Do not overcook the OFSP**

Scoring

Each step was allocated a score of 1 if performed well. The maximum score that any mother could get was 6.

4.6 Memorisation/recall

Did the mothers managed to retrieve the stored health messages?[77]

Once a health message is understood, people must remember the message in order to use it.

Memory is probably the most widely studied aspect of cognition as it is implicated in all of the cognitive tasks we perform. For example: in order to read, write, speak and understand spoken language we have to access information about word meaning and grammatical structure that are stored in our heads. In order to drive a car or ride a bike we have to access procedural knowledge about how to perform these actions. We also rely on detailed geographical knowledge about routes to work and the layout of the town where you live. [139]

4.6.1 What is Memory?

"It is the retention of information over time. Without memory, one would not be able to connect what happened in the past to what is happening today. Educational psychologists study how information is initially placed or encoded into memory, how it is retained or stored after being encoded, and how it is found or retrieved for a certain purpose later." [67] It is these three important steps – encoding, storage and retrieval that make up memory.

Encoding: When a student is listening to a teacher, he/she is encoding information into memory. Encoding will only occur if the student pays attention. [67] This shows that there is a directional link between attention and memory. Research has shown that deep processing, elaboration and repetition of information over time improves encoding for long-term memory. This means giving meaning aided with examples to the information and also constructing images for that example. After this, the information needs to be organized. [67, 154]

Storage: after encoding the information, students need to store that information. A simplified model is that of the Multi-store model of memory which considers the structure of human memory into 3 main stores: sensory memory, short-term memory and long-term memory. Students are exposed to sensory information for a brief seconds, and that information to which they pay attention is then transferred to short-term memory which is also limited-capacity but lasts longer than sensory memory (i.e. about 30 seconds). However, if the information is repeated over time it can be retained

longer, hence increasing the chances of it being transferred to long-term memory. The latter can hold enormous amounts of information for a long period of time in a relatively permanent fashion. Information in long-term memory can then be retrieved back into short-term memory. [155] However, the multi-store model of memory has several weaknesses. For example, long-term memory is treated as a unitary structure when it seems clear that there are many types of long-term memory. Researchers often distinguish between declarative and procedural memory, where declarative memory holds factual information and procedural memory is memory for sequences of action (such as knowing how to ride a bike). [154]

Retrieval/ recall: Long-term storage will be of no good unless we can retrieve stored information when we need to use it. When we retrieve something from our mental “data bank”, we search our store of memory to find the relevant information. [67]

Memory can be forgot and one source of forgetting is the passage of time that leads to memory decay. [67, 154]

Short-term memory seems to be dependent on attention. Long-term memory seems to be dependent on attention and understanding for good encoding and also, there is a relation between attitudes and memory in that, attitude duration will depend on memorisation of that attitude. [77]

4.6.2 Measures of memory – recall:

Many studies of human memory have been laboratory-based and involve presenting subjects with sets of words, letters or numbers and asking them to recall or recognise these words at a later point in time. During a recall task the subject is asked to write down as many of the test items as they can remember. Psychologists have developed detailed accounts of human memory based on the data from these types of laboratory task. Digit or letter span tasks provide evidence for separate short and long-term memory systems. In these experiments, subjects are presented with a list of items (usually numbers or letters) for subsequent recall. A standard finding is that subjects are good at recalling items that appeared early in the list (these items have been stored in long-term memory) and items that appeared late in the list (these items resided in short-term memory at the time of testing), but have poor recall for items that appeared in the middle of the list (these were not encoded in long-term before being displaced from short-term memory. Another finding is that subjects retrieve information better

when using cues. These cues could be information that has been elaborated or images about the item that one wants to retrieve. [139, 156]

Recall can be measured as cued recall or free recall, and success in them is based upon the subject's knowledge of events that occurred when he/she was personally present in a particular spatio-temporal context. In free recall subjects are asked to repeat what they heard or read without any cues or prompts. In cued recall, information is first presented in conjunction with some other stimulus and, when testing recall the other stimulus is presented as a cue to stimulate recall. In the context of health education, free recall occurs when a patient reads or hears information about a health problem and, later, without any pictures or cues, remembers that information in deciding what actions to take or to tell to other people. Cued recall occurs when a patient reads or hears health information with an accompanying picture and later views the same picture to help remember the information.[157] Literature demonstrates that pictorial presentations are superior to verbal presentations when it comes to free and cued recall.[76]

4.6.3 Measuring recall to health messages:

The following studies give an example of how recall to health messages has been measured.

Of the studies that have attempted to measure attention to health messages, they have done so by :

1. Asking subjects to read instructions (for example for preparing and administering a solution for the treatment of dehydration in children) and then asking them to recall later. Correct answers are then scored. [81] [158]
2. Asking subjects to recall the content of health education material (for example magazines or posters). Respondents could also be given the education material, but masked and then asked to recall the missing information. Again, recall of correct content is scored and can be transformed to percentages.[87] [88] [131] [134]
3. Comparing the number of total and key messages in the content-analysed materials with the number of total and key messages that were retained and recalled by the respondents in the focus groups. [84]

In Rimmer and Glassman's (1984) study, memorisation was called retention, and it was defined as the degree of comprehension at a point four weeks later in time. Both comprehension and retention were measured in the same way, only the time interval

differed between the two. Subject's answers were scored and then compared at the 4-week period. [78]

4.6.4 Methods used to measure recall in the current research:

Recall levels of the mothers was measured instead of the McGuire's Memory. Because memory can decay, it is important to measure recall capacity after some time and not straight after the intervention like some studies do. Hence, in the current study recall was a measure of long-term memory since at least 2 months passed since the time that the mother was exposed to the message until the time that she was asked to retrieve the information.

Comprehension Lexical meaning, could also have been considered as a measure of short-term recall. Likewise, psychological-operational comprehension assessment could have been considered as a measure of recall – long term memory. This was the case for Rimmer study as she measures both constructs using the same questions.

Recall in this study was measured immediately after participants had carried out the recipe task (i.e. the skills acquisition task). (*Tool can be seen in appendix 4.3*)

The first recall question measured subject's ability to recall correctly the ingredients used in their chosen recipe.

The second recall question, measured subject's ability to identify the right food group for each of the ingredients used in their chosen recipe. To aid recall (hence, "cued recall"), the health promoter showed a outlined pencil-drawing of the 4 food groups, but emptied circles (i.e. they were used to see the drawings of the foods inside each food group). Those mothers that chose the recipe with 4 ingredients received 4 of these questions, while those mothers that chose the recipe with 2 ingredients only received 2 of these questions.

The third recall question, measured subject's ability to identify the number of recipes that were carried out in the demonstration.

The fourth recall question, measured subject's ability to identify all the ingredients in each recipe.

Scoring

The first item was awarded a point if correct. If there was one ingredient missing or one mentioned that was not part of the recipe then the subject did not receive the point.

The second item was awarded a point for each correct answer. However, total score depended on the number of ingredient used. Scores could range from 2-4, therefore a proportion of the total possible scores was calculated.

The third item was awarded a point if correct. And this was based on the recall of the correct number of recipes demonstrated.

The fourth item was awarded a point for each correct recall of ingredients in each recipe.

All of the scores in the items were summed up and a percentage of the total scores was allocated for each subject.

Piloting/pre-test

Given that this task was complex for the setting of the study and involved different logistic, it was pre-tested before hand on 5 mothers who were then not included in the actual measurements. The aim of the pre-test was to measure the clarity and ambiguity of the comprehension and recall questions and check the respective answers and also the feasibility and practicality of the skills acquisition task. Originally there were 8 steps for the skills acquisition task, but two were dropped at it would lead to differences in the scoring between mothers. It is possible that perhaps 5 people was a too small number for the pre-test and maybe more mothers should have been included in this.

4.7 Attitude change

Did subjects improve their attitudes regarding breastfeeding, dietary and feeding practices?

This step like all the others that precede it was measured cross-sectionally. This experimental design constraint precluded the use of the "change in attitude" step. Hence, "positive attitude" alone was used. The conceptual definition for the term was the receiver's attitude toward the message at the point where the measurement was

taken. However, to account for the lack of a baseline measurement which constraints the measurement of change, the TSNi control group was included. The control group was here used as a proxy for baseline. In many studies, there is no danger of doing so, however, in a 2.5 year study there is the risk that the control group might have also received some nutrition education information from other sources and it was the case that the control group was contaminated from also being exposed to the radio program which also disseminated messages on the promotion of OFSP, some breastfeeding practices and some complementary feeding practices.

4.7.1 The nature of attitudes

Despite the long history of research on attitudes, there is no universally agreed-upon definition.

Fishbein and Ajzen (1975)[159] define attitudes as “learned predispositions to respond in a consistently favourable or unfavourable way towards a given object, person, or event.” This definition emphasizes three features of attitudes: firstly, they are learned; secondly, they are consistent; and thirdly they are concerned with favourable or unfavourable responses.[159]

There is a consensus that at the core of all the definitions is the notion of evaluation. Thus “attitudes are commonly viewed as a summary evaluations of objects (e.g. oneself, other people, issues, etc) along a dimension ranging from positive to negative”. [160] [161]

The use of attitudinal measures have become important for a long time as sociologists have been assuming that attitudes correlate substantially with behaviours, so, there was this assumption that individuals tend to act in accordance with their attitudes. As can be seen from McGuire’s sequence of output steps, attitude almost precedes behaviour change.[162, 163] Apart from the possible attitude-behaviour relationship, attitudes are an important concept that help people understand their social world by guiding people’s reactions towards objects and features in our environment. [164, 165]

4.7.2 Attitude formation and change

People are always adopting, modifying, and relinquishing attitudes to fit their ever changing needs and interests.

Most psychologists agree that attitudes can be acquired from others (i.e. social learning) in the form of classical conditioning and instrumental conditioning, and modelling. One's interaction with others, such as parents, teachers, peers, relatives, newspapers, books, television, religious groups, etc., can effect our attitudes toward various objects. This learning comes from situations in which we interact with them or merely by observing their behaviours, but also through direct personal experience. [159, 164-166].

There have been many theoretical approaches to attitude change. Among these, many researchers have looked at the relationship between persuasive communication and attitude change. They believe that attitude cannot be changed by simple education. Acceptance of new attitudes depends on who is representing the knowledge, how it is presented how the person is perceived, the credibility of the communicator, and the conditions by which the knowledge was received. [165]. Again, a change in attitude will be more effective if the persuasive message takes into account the source of the message, message variables, and receivers of messages. [165]

The elaboration Likelihood model and the Heuristic model of persuasion are the main models looking at the relationship between persuasion and attitude change.

4.7.3 The attitude-behaviour relationship

A question that has persistently challenged investigators interested in the relationships between attitudes and behaviour is: Do attitudes predict and/or cause behaviour?

There have been two schools of thoughts, one belonging to a group that believes attitudes to predict and cause behaviour, and another that instead, believe that it is the experienced behaviour that determines attitudes. [162]

Ajzen and Fishbein argue that this lack or weak relationship has been due to the methodology used to assess this relationship which has been weak. They developed the Theory of Planned Behaviour where they argue that an assessment of behavioural intentions will lead to the most accurate behavioural predictions, but if, attitudes and behaviour of comparable levels of specificity are measured, even attitudes will be found to be excellent predictors of behaviour. Ajzen and Fishbein have thus modified the Attitude-Behaviour relationship to an Attitude- intention-behaviour relationship. [162, 167] This theory agrees with McGuire's where decision-making is the equivalent of intention.

The second school of thought, argues that, taken alone, attitudes are often inadequate predictors of behaviours, that other mediational processes and variables such as experience, perception, cognition, context and goals also have a role to play. [162] The most widely cited review has been that of Fazio and Zanna [162] which indicated that attitudes based on direct experience with the attitude object are typically more predictive of behaviours than attitudes not so based. [162] In a meta-analysis, Wallace *et al* (2005)[168] showed that situational constraints, such as perceived social pressure and perceived difficulty, weaken the relationship between attitudes and behaviour. [168]

4.7.4 Development of the attitude instrument in the current research study

The attitude measurements were carried out for approximately one month and a half, from end of October 2004 until mid December 2004, after the nutrition education program had terminated. Attitudes were measured at the same time as the TSNI end-of-project Survey. While subjects were waiting to be interviewed for the TSNI end-of-project Survey they were selected at random to be measured for attitudes. This had to be well coordinated since both the Survey and the attitude instrument were measuring the same subjects.

Overall, the attitude tool was administered to 139 subjects from intervention I, 142 from intervention II and 130 from the control group.

The researcher supervised all the measurements, but because the tool had to be administered in the local language, someone else fluent in the language was contracted and trained to do it.

4.7.4.1 Mode of administration

An attitude tool usually asks/probes people's agreement towards a subject/object/action. The probing/asking can be carried out using a questionnaire, and the questionnaire itself could be made out of questions (for example: "Do you agree that mothers should breastfeed exclusively until the 6th month?") or statements (for example: "It is important that mothers breastfeed exclusively in the first 6 months."). In the current research study a story was presented to the mothers who were then asked

for their agreement/disagreement towards the story. The advantage of using a story is that it is less direct and invasive, and the respondent does not feel as if she is giving an answer about her personal life, while she is still expressing her feelings and beliefs towards the attitude object/action/subject.[169] (See *appendix 4.6*)

Along with the story, mothers were also presented with vignettes depicting the story. A review by Houts[76] concluded that adding pictures to spoken language can increase attention and comprehension. Again, the advantage of using vignette is that there is a possibility that the question might make the respondent feel uncomfortable, they might feel that they are being judged by their replies, and by using vignettes about other people they do not feel the same threat, especially if the scenarios are believable. But has shown by Houts, vignettes make the stories easy to understand and to some extent real. Most vignettes were pictures taken locally, hence making the stories more real and representative of the local population. (See *appendix 4.6*)

While it is not usual to ask the beliefs/reasons behind the subject's attitude, in the current study approximately 60% of the total sample in each intervention group was asked to elaborate the reasons for their attitudes. The main reason for this was not only to see if the scaling response (using the Likert scale) matched the answers thus checking that subjects understood the story but also, to gain more insight on the beliefs behind the attitudes.

4.7.4.2 Devising the items

The TSNI nutrition education program aimed at improving breastfeeding, and child feeding practices. Hence, it would only make sense that the attitude instrument developed for the current research should address these topics.

The finalised attitude tool comprised of two distinct sections, section A referring to breastfeeding practices and section B referring to complementary feeding practices. (See *appendix 4.4*)

Seven of the total items on section A referred to strongly rooted traditional practices related to taboos, the type that would need a lot of persuasion to show any positive change. One item was used as a control, and the other was chosen to reflect the importance that the mother would give to breastfeeding. This information was based on results from the TSNI baseline survey and other previous research. The latter item was

the only one that was not mentioned in the classes, while all the other items were part of the persuasive communication curriculum of the group classes.

Most of the items on section B, except item 2 which refers to a more traditional practice, referred to incorrect feeding practices. The choice of items was a representation of results from TSNI baseline survey, other previous research, but mainly from a formative research for the TSNI study, TIPs, which included home observations on 12 households. Apart from items 2, 4 and 9, all the other items were part of the 12 key recommendations given repeatedly at IC. They were also part of the persuasive communication curriculum of the group classes even though not as frequently mentioned. Items 2 and 4 were part of recipe demonstrations, which was carried out in GC. And item 9 was chosen to reflect the importance that the mother gives in the quality of the child's diet.

A review of the literature was unable to find any studies measuring attitude regarding breastfeeding and complementary feeding practices in Mozambique (or in another Sub-Saharan country). Hence, in the current research study, items for the attitude instrument had to be developed *de novo*.

In each of the sections the items were related to each other, hence, the items was designed as a scale, however, items were individually compared between groups as batteries. [116]

The stories-vignettes for each item were based on research carried out on local practices, and most of the vignettes were in fact pictures taken locally. Overall, 18 items were chosen after being piloted. The following is a description of each item. An example of the attitude tool administered in the field, the vignettes and the Likert scale tool can be seen in Appendix 4.4.

Section A –Breastfeeding practices

1st item- *Is colostrum good?*: results from the TSNI baseline survey showed that the majority of mothers rejected the colostrum, as it was seen as a dirty and disease-causing liquid. The accompanying story-vignette told/showed a grandmother squeezing the breast of her daughter 3 days after she had given birth so that the first milk, that has a yellowish colour could be discarded.

2nd item- *Are traditional medicines good for the infant?*: previous research showed that giving a traditional tea to a newborn was widely practiced. Local communities believed

that this tea is protective of a milk-causing disease. The first accompanying story-vignette told/showed the same woman that gave birth going to a traditional healer to acquire leaves and roots so that she could prepare a tea for the baby. The second accompanying story-vignette told/showed this woman giving this tea to her newborn baby.

3rd item- *Is breastfeeding good for diarrhoea?*: results from the TSNI baseline survey showed that the majority of mothers thought it was the milk that was causing the diarrhoea and that therefore the mother should no longer breastfeed. The accompanying story-vignette told/showed the baby 3 months old with diarrhoea. The mother noticed that that if she breastfed less often the baby would also have less diarrhoea, therefore she decided to decrease the frequency of breastfeeding.

4th item- *Is breastfeeding the child more important than working in the field?*: this was not based on any previous research. It was chosen to give an idea of the importance that mothers infer to breast feeding a small child. The accompanying story-vignette told/showed explains that it is the month of rains, and that the mother of the baby is taking the most of the rain to work in the fields. However she is becoming tired of caring her 6 month child in her back. Because of the age of the child the mother decided to stop breast feeding the child so that she could leave the child at home with the grandmother.

5th item- *Is it acceptable for a 1-month pregnant mother to breastfeed the child?*: results from the TSNI baseline survey showed that the majority of mothers thought that pregnant mothers should not breastfeed their child. However, the story and vignette showed a woman in her early pregnancy where the belly was still not noticeable.

6th item- *Is it acceptable for a 4-month pregnant mother to breastfeed the child?*: in this story-vignette the same pregnant woman already has a noticeable belly.

7th item- *Is it okay for a sick mother (with malaria) to continue breastfeed her (3-month) child?* Results from the TSNI baseline survey showed that the majority of mothers thought that sick mothers should not breastfeed their child for fear of contaminating the child. However, the question in the TSNI survey did not specify the age of the child. A group of women were consulted regarding this practice and these explained that the child would be too small to stop breastfeeding. This story-vignette tells of a mother that gets sick with malaria and decided to stop breastfeeding her 3-month child, however,

her sister advises to continue breastfeeding which the former does. This item was used as control as it would be expected for subjects to agree with the story-vignette.

8th item- *Is it okay for a sick mother (with malaria) to continue breastfeeding her child (of 9 months)?* This story-vignette is the same as above, but the child is grown and the sister is not present. Thus, the sick mother decides to stop breast-feeding.

9th item – *Is It okay for a mother to continue breastfeeding her child after an interruption?* Results from the TSNI baseline survey showed that the majority of mothers thought that breastfeeding should stop if the mother has discontinued breastfeeding for some time. The belief is that the breast milk has become rotten and there us a risk of contaminating the child. This story-vignette tells of a mother that had to travel for 3 days and left her then 9-month breastfeeding baby with her sister. In her return she decided not to breastfeed the child again.

Section B –Complementary feeding practices

10th item- *Are two meals a day enough for a crawling baby?* Results from the TSNI baseline survey showed that the majority of mothers thought it was enough for a crawling baby to eat only twice a day. The story-vignette tells of a crawling baby that was fed twice a day because his/her mother thought it enough.

11th item- *Should an egg be added to the porridge of a 6-month old child?* Previous research showed that women thought that a baby should not eat egg as he/she was still very young and also as it my lead to rotten teeth. The story-vignette shows a mother preparing a porridge for her 6-month infant and deciding not to add an egg as she thinks it is not appropriate for a baby of that age.

12th item- *Should a mother motivate her child when he/she refuses to eat?* Household observations and TIPs showed that many times the mothers do not stimulate their children when they refuse to eat. The story-vignette shows a mother with her child in her lap trying to feed him and the child turning his head away. The mother would do nothing and would just wait until the child was hungry again.

13th item- *What is a good porridge for the child?* The vignette- story explains that for that mother the best porridge for the child is one consisting of cassava flour and sugar as it fills the child's belly. Household observations showed that the majority of caregivers prepare weaning foods like this one. Because the nutrition education

included a lot of teaching and practices recipes on enriched porridges, it was important to measure the perceptions of what should be a good porridge for the child.

14th item- *Is it okay for a 2-yr old child to skip breakfast and only have lunch and dinner?* Results from the TSNi baseline survey showed that knowledge about the frequency of eating of a child at different ages was very poor. The story-vignette tells of a mother that needs to go to the field very early in the morning and therefore her child does not have breakfast. But the mother is not very worried about it because the child already walks and therefore having lunch and dinner is enough.

15th item- *Is it enough for a child to eat fat rich foods 3 times a week?* TIPS showed that children were eating very few sources of fat. Groups classes but especially individual sessions mentioned that the child should eat fat sources everyday and if possible 3 times a day. The story-vignette showed the sources of fat available locally and explained that the mother of the child in the picture would give these foods 3 times a week because it was enough.

16th item- *Is it appropriate for a 2-year old child to eat without any supervision?* Household observations showed that many mothers leave their young children unattended while they are eating. The story-vignette tells of a mother that leaves the child in one part of the house to eat without supervision while she goes to another part of the house.

17th item- *Is it appropriate for a child to share the plate with her parents?* Household observations and TIPS showed that it was extremely normal for a small child to share the plate with parents or siblings. The story-vignette tells of a child that his sharing the breakfast in the same plate as her parents. This does not worry the mother as she thinks that the child will eat the same amount as if she was eating on her own.

18th item- *Is it appropriate for the father to eat liver and not the rest of the family?* This story-vignette tells of family that is having a dish made of dark green leaves and coconut for dinner. However, the father is also having liver, and the reason why no one else is having it, including the child, is because there was not enough for everyone.

Since 61% of the women recruited for the TSNi study had no formal education, the attitude tool had to be extremely simple and clear without employing any technical terms. Negative wording of the items was avoided, however, the majority of items were negatively-keyed, meaning that the items were phrased so that an agreement with the

item represented a relatively low level of the attribute being measured, or, a negative attitude.[116] One of the few items showing a correct practice was included as a control. So if most of the respondents disagreed with this control item, then something would probably be wrong with the tool or with the subject understanding of that item. Also, the person carrying out the measurements was careful to detect those subjects that tended to agree or strongly agree with all the items. If this were the case then a very absurd question would be asked to test the subject. All those cases where the subject (strongly) agreed with all the items, which always agreed as well with the absurd question, were not included in the analysis.

It is important that the attitude scale does not include too many items in order to avoid subject burden. However, in order to try to cover most of topics taught in the class, 9 items were taught sufficient for each topic without causing too much subject burden.

The assistant and one of the health promoters translated all items to the local language, Chuabo. After translation, the translated items were tested on the remaining 3 health promoters in order to check if the meaning was still the same as in Portuguese. This not only included the attitude question but the story, which went along with the vignette (see appendix 4.6). To facilitate the work of the assistant, which was responsible for administration of the tool, all vignette cards included in their back, the translated story and question.

4.7.4.3 Scale

Likert scales was the scaling method chosen for the attitude instrument developed in the current research. A 5-point Likert scale was used in favour of a 3-point or a more extended Likert scale. An extended (i.e. 7-point) Likert scale might have confused the subjects, and a 3-point Likert scale, while simpler, would not allow for levels of agreement/disagreement. The advantage of allowing for the use of the extreme (i.e. 1 and 5) and the middle points (i.e. 2 and 4) in a 5-point Likert scale is that it reflects some careful thought and weighing of answers. [116, 169]

After being presented with the story-vignette, the mothers were asked whether they agreed, disagreed or were not sure with the practice in the story. Depending on their answer they were then given to chose the levels of their agreement/disagreement. However, to make it easier for them, they were presented with faces that corresponded to these levels of agreement/disagreement (see appendix 4.5). An example would be: *"During the first days after baby Maria was born, Luisa's mother was squeezing her*

breast in order to discard of that first yellow milk." This story would be told along with a vignette depicting it. Then the mother would be asked: "Do you agree, disagree or are indecisive with Luisa's decision to discard this milk?" If the mother answered for example that she agreed, then a photo of a man agreeing and another of him agreeing very much was presented to the mother and the difference between the two explained. The mother then had to pick the face that best represented her level of agreement. The face chosen belonged to one of the enumerators which was someone already known in the study areas, liked by everyone due to his constant cheerfulness. During the pre-test, mothers' reactions to the male faces were tested and the conclusion was that there were no culture obstacles and in fact, the respondents found it funny.

Scoring

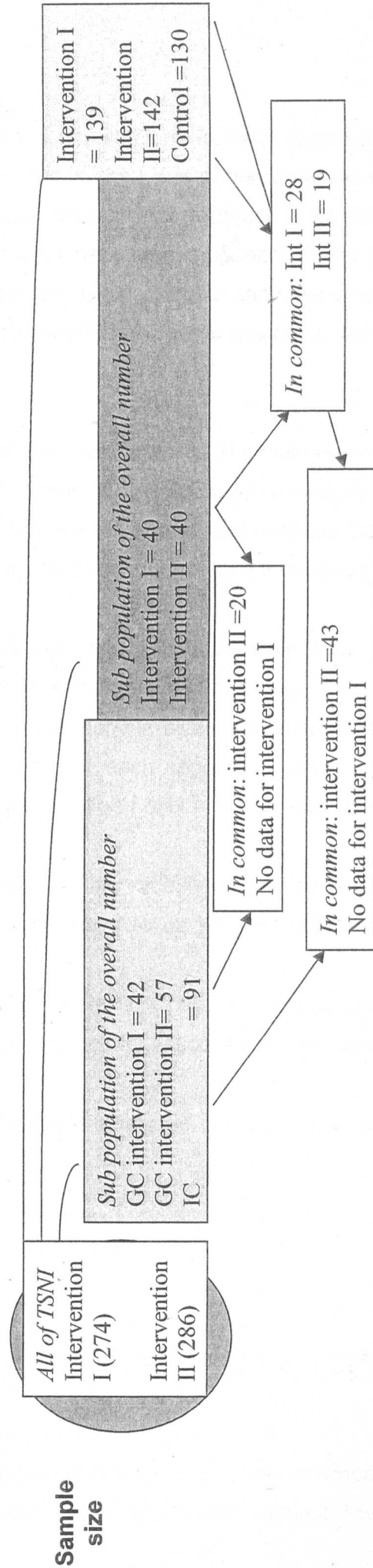
Each section comprised of 9 items and was scored independently of the other. Because the attitude tool includes positively-keyed and negatively-keyed items, the negatively-keyed items must be "reverse-scored" (i.e. 1=5, 2=4, 4=2, 5=1). This ensures that all of the items – those that are originally negatively-keyed and those that are positively-keyed – are consistent with each other, in terms of what an "agree" or "disagree" imply. Hence, total attitude scores in each section could range from 9 (indicating negative attitudes towards practices) to 45 (reflecting positive attitudes towards practices). A score of 27 indicated a neutral attitude.

Piloting

A pilot test was carried out on 10 women in order to test responses, specifically to test the clarity, understanding and ambiguity of the attitude items and vignettes. Ambiguous and unclear items were removed or improved. These women measured for the pilot test were not included in the final measurement.

Table 4.1 Summary of methods and sample size for each output step

Outcome/output step						
	Exposure	Attention observation	Attention recall	Comprehension-Lexical meaning	Comprehension-Psychological meaning	Skills acquisition
Method used for the measurement	Block 2 output steps					
	Attendance rates [proportion of attendance to overall classes]	Attention scores [number of observations (out of 10) where the mother was paying attention]	Number of questions (out of 2) answered correctly	Was question answered correctly (No/yes)	Comprehension scores [proportion of questions answered correctly – score of 1 for each correct answer]	Skills scores [number of steps (out of 6) which the mother performed well]
						Recall scores [proportion of questions answered correctly – score of 1 for each correct answer]
Timeline for the measurement	Block 3 output steps					
	Attitude questionnaire rated with a 5-Likert scale	Recall scores	Skills scores	Recall scores	Attitude questionnaire rated with a 5-Likert scale	Composed of two independent sections each consisting of 9 items
Timeline for the measurement		May 2003 – November 2004	May 2004-September 2004	May 2004-September 2004	Recipe task: October-November 2004	Recipe task: October-November 2004
Sample size		Intervention I (274) Intervention II (286)	May 2004-September 2004	May 2004-September 2004	Recipe task: October-November 2004	Recipe task: October-November 2004
Sample size		Intervention I = 139 Intervention II = 142 Control = 130	Recipe task: October-November 2004	Recipe task: October-November 2004	Recipe task: October-November 2004	Recipe task: October-November 2004



4.8. Data analyses

This analysis is based on the assumption that the observations in each approach-intervention group are independent of each other, that is, the value of one observation is not influenced by the value of another. This assumption of independence is violated if the data are clustered, that is, if observations in one cluster tend to be more similar to each other than do individuals in the rest of the sample. Clustered data can arise when subjects are grouped.^[170-173] In the case of this research study, some outcomes were measured in a group class setting.

When possible, clustering should be allowed for in statistical analysis. The main reason for this is that standard errors may be too small if they do not take account of clustering in the data. This will lead to confidence intervals that are too narrow and p-values that are too small. [170-173] However, data analysis in this study did not take into account clustering for several reasons:

- Clustering was not accounted for in sample size calculations, because it would have resulted in a logistically impossibly large sample size for each output step;
- Since clustering was not accounted for in the sample size calculations and during design of the study, the number of clusters for each approach-intervention group was very small. For example, block 2 GC intervention I only had 5 clusters and block 3 intervention I and II had 6 clusters each;
- Clustering in IC would mean that there would be a big variation in the size of the clusters and therefore, the very small clusters would have to be joined together to equal the size of the bigger clusters;
- If the number of clusters is very small, then there is a reduction in possible variation and therefore it becomes almost impossible to look for associations with other variables;
- If some output steps have the cluster effect and others do not, then it is not possible to correlate them.

4.8.1 Measuring the distribution

Histograms were plotted for each output step by approach-intervention group in order to have an idea of the distribution of the data.

P-P plots were used to determine whether the distribution of each output step matched a normal distribution. They plot a variable's cumulative proportions against the

cumulative proportions of any of a number of test distributions. If the selected variable matches the test distribution, the points cluster around a straight line. [170, 174, 175]

4.8.2 Tests for the comparison of groups

Exposure/attendance levels, attention observation, attention-recall, and comprehension lexical meaning, compared three groups: GC intervention I, GC intervention II and IC. Analysis of Variance, which included the Bonferroni correction, was used to compare the three groups when the output step was continuous. Because GC consisted of 12 classes and IC only 6, a proportion of attendance was instead used when comparing the 3 groups for the exposure output step. For attention recall and comprehension lexical meaning output steps which are binary variables, a Pearson's Chi-square test was instead used to compare the groups. Moreover, intervention groups were compared between each other using an Independent T-test. For the exposure output step, GC intervention II and IC mothers were the same, therefore their proportion of attendance to both GC and IC were summed and then averaged to give a proportion of attendance for the intervention II group, which was then compared with that one for intervention I.

Comprehension psychological meaning, skills acquisition and recall, compared two groups: GC intervention I and IC. A T-test, was used to compare the two groups.

Attitude change compared three groups: Intervention I, Intervention II and a control group. Analyses comprised of a secondary and a primary analyses. The secondary analyses consisted of comparing each item between the 3 groups. For this, points 1 and 2 were aggregated, so were 4 and 5. Since this was carried out before reverse scoring, 1 and 2 reflected a positive attitude (except for items 5 and 7), and 4 and 5 a negative attitude. A Chi-square analysis and risk (odds) ratio was used for this. The primary analyses consisted of comparing the summated scores for each section between the 3 groups. Analysis of Variance, which included the Bonferroni correction, was used for this.

Confounding was accounted for whenever necessary. Multiple regression was used if the main outcome (i.e. results for output steps) was continuous, and logistic regression was used instead if the main outcome was binary.

4.8.3 Tests for correlations and associations

Correlations and associations between output steps and factors/exposures

These were carried out separately for each approach-intervention group.

Associations can be measured using correlation tests. However, correlation analysis do not inform us of the predictive power of variables. For this reason, regression analysis would be needed since it fits a predictive model to the data and then uses that model to predict values of the dependent variable from one or more independent variables. [170, 174, 175]

It is difficult to decide what is the appropriate test of association to use since there is a range of choice. The following is a description of the possible ones:

- **Pearson's r** is the most widely-used type of correlation coefficient, also called linear or product- moment correlation.

- **Spearman's rho** is the most widely-used type of correlation coefficient when the distribution is not normal. As a test for association can examine how a numerical (i.e. discrete or continuous variable) output step varies according to a categorical variable. [170, 174, 175]

- **T-test** would be useful to compare a binary variable against a continuous one, when the Pearson's r is not the best test. For every T-test result, Levene's test result was used to decide whether to read equal or unequal variances' p-values.

- **Linear regression** which estimates the best-fitting straight line to describe an association between a numerical outcome and a numerical exposure. This measure is only meaningful if one wants to create a regression equation. [170, 174, 175]

- **Chi-square test for 2x2 and larger contingency tables:** it examines the relationship/association between two categorical variables. [170, 174, 175]

The following matrix (table 4.2) shows a number of possible tests for the associations between the main outcomes (i.e. output steps) and the exposure. However, only those tests written in bold were the ones used for the association analysis. The justification for the choice of these test derived from the classification of the variables (which is written in italics) into their different types and the reason behind the use of test. So for example, Pearson's r was used instead of Linear Regression, because there is not a need to find out the regression equation for the association between the two variables.

When looking at correlations a considerable amount of caution must be taken when interpreting correlation coefficients because they give no indication of the direction of causality. In any bivariate correlation causality between two variables cannot be assumed because there may be other measured or unmeasured variables affecting the results. Correlation coefficient say nothing about which variable causes the other to change. Sometimes, the researcher has to use of intuition and logic to assume which variable is having an effect on the other. [170, 174, 175] This is the case for this research study.

It is the size of the correlation coefficient and not the p value that dictates the strength of the association between two variables. Cohen [176] (1988) for example, has suggested the following interpretations for correlation coefficients: weak – 0.10 to 0.29; medium -- 0.30 to 0.49; and strong – 0.50 to 1.00. This study will be using Cohen's interpretations along with p-values. All $r < 0.30$ will be interpreted as no association.

Associations between output steps

Association between the output steps was also measured, simply to test McGuire's theory that they are all linked.

The majority of associations was carried out using Pearson's r which measures the strength between the two output steps. For binary output steps, a T-test was instead used.

It is important to note that sample size differed for the different blocks of output steps. So while exposure was measured for the whole TSNI sample size and therefore shared the same sample size as all the other output steps (except with GC intervention I in block 2 output steps as their id was not known), block 2 output steps (i.e. attention observation, attention recall and comprehension-lexical meaning) only shared 20 IC mothers with block 3 output steps (i.e. comprehension-psychological meaning, skills acquisition, and recall) and 43 intervention II mothers with the attitude output step. Block 3 output steps only shared 25 intervention I mothers and 22 intervention II mothers with the attitude output step.

Cohen's interpretations for the size of r will also be used to interpret these associations.

Reliability analysis

Test-retest reliability was calculated for the attitude questionnaire and for attention recall and attention observations.

A sample size of 10 mothers was chosen to carry out this type of reliability analysis. With this approach the reliability coefficient is obtained by simply calculating the correlation, using the Pearson product moment correlation, between the scores on two administrations. Because reliability varies over time, with the attitude tool, 4-7 days were given between the two measurements. The risk of having a shorter time period between administrations is that subjects might remember some of the questions and their responses. A minimum coefficient of 0.70 is recommended in research.[116] [117]

Internal consistency reliability was also measured for the attitude questionnaire using Cronbach's alpha. This judges the reliability of the instrument by estimating how well the items that reflect the same construct yield similar results. It looks at how consistent the results are for the different items or the same construct within the measure. Only one administration of the test is needed, and hence, scores for all 411 subjects (i.e. mothers in the 3 groups) were used. The test was carried out separately for both sections of the tool.[116] [117] Alpha coefficient ranges in value from 0 to 1. The higher the score, the more reliable the generated scale is. Nunnally (1978) has indicated 0.70 to be an acceptable reliability coefficient but lower thresholds are sometimes used in the literature. [177]

4.8.4 Missing values

Some exposure factors which were measured during the measurement of the output steps (ex: disturbance levels), were not measured for all mothers. Hence, associations with these factors were not carried out for all mothers.

The same thing happened with the socio-demographic variables. The TSNi surveys did not collect information for all households, and for that reason, the associations were not carried out for the whole sample size. This was the case with the education level of the father whereby data was not collected for 37 fathers in intervention I and 44 in intervention II, and also with income per capita whereby data was not collected for 28 intervention I households and 34 intervention II households, additionally, mother's perception data was not collected for 20 mothers in intervention I and 28 in intervention II.

Table 4.2 Tests for association between output steps and exposure variables

Output steps:	Exposure	Attention Observation	Attention recall ¹	Comprehension-lexical meaning	Comprehension Psychological meaning	Skills acquisition	Recall	Attitudes
Factors/exposure:	Continuous	Continuous	Binary	Binary	Continuous	Continuous	Continuous	Continuous
Age (years)	Linear regression	Linear regression	Pearson's r	T-test	Linear regression	Linear regression	Linear regression	Linear regression
Continuous	Pearson's r	Pearson's r	T-test		Pearson's r	Pearson's r	Pearson's r	Pearson's r
HH size	Linear regression							
Continuous	Pearson's r							
Education level of mother and father (schooling years)	Linear regression	Linear regression	T-test	T-test	Linear regression	Linear regression	Linear regression	Linear regression
Continuous, but also binary (i.e. has/ does not have formal education)	Pearson's r	Pearson's r	X ² test	X ² test	Pearson's r	Parson's r	Pearson's r	Pearson's r
	T-test	T-test			T-test	T-test	T-test	T-test
Per capita income (MZM)	Linear regression	Linear regression	T-test	T-test	Linear regression	Linear regression	Linear regression	Linear regression
Continuous	Pearson's r	Pearson's r			Pearson's r	Pearson's r	Pearson's r	Pearson's r
Mother's perception of child's health	T-test							
Binary								
Disturbance levels		Linear regression	T-test	T-test				
Continuous		Pearson's r						

Class size <i>Continuous</i>	Linear regression Pearson's r	T-test	T-test				
Duration of class (minutes) <i>Continuous</i>	Linear regression Pearson's r	T-test	T-test				
Time interval (days) <i>Continuous</i>				Linear regression Pearson's r	Linear regression Pearson's r	Linear regression Pearson's r	
Cooking method (vapour/boiling) <i>Binary</i>				T-test	T-test	T-test	
Type of recipe (1,2,3,4 or 5) <i>Categorical</i>				Spearman's Rho	Spearman's Rho	Spearman's Rho	
Number of ingredients (2,3 or 4) <i>Categorical</i>				Spearman's Rho	Spearman's Rho	Spearman's Rho	

Note: some variables that were collected as numerical could have also been categorised

¹ Attention recall was categorised into 0 – no correct responses, 1 correct response, and 2 correct responses. However, only 1 mother answered incorrectly both questions. This value was included with the other values of mother that got only 1 correct answer, and instead attention recall was treated as a binary variable.

4.8.5 Measure of persuasiveness

All the previous tests were used to compare results of each output step between the intervention groups. While this might tell us how one intervention group compares against the other for each output step, it does not give us a overall picture of the degree of persuasiveness of the persuasive communication (i.e. nutrition education) delivered with one intervention against the other. Following McGuire's "sequence of output steps, a high persuasiveness degree would imply that the person has been exposed, paid attention, understood, agreed with the message(s), gained skills and later, was able to recall the message(s).

This variable, which shall be called persuasiveness degree, is in fact a weighted composition of all the output steps. As McGuire argues, a person would have to go through all the steps to be persuaded, persuasion being changing behaviour. He also mentions that earlier steps in the model-sequence have a weaker impact in the last steps. For this reason, weights were allocated for each output step, whereby the earlier output steps (ex: exposure and attention) received lower weights than those steps found more towards the end of the sequence (ex: recall and skills acquisition).

This approach is also used in health measurement scales, whereby a number of items are used to measure an underlying characteristic. A simplistic way would be to add all the items but this assumes that all the items are equally important. Rather than doing this, a scale or index can be developed which weights each item.[116]

An initial approach considered for this research study was to calculate the weights by regressing each output step with behaviour change and dividing the r value by the standard deviation of each independent output step. The reason for this comes from the fact that each output step used different measurement scales, hence by dividing by their SD we would be standardising the results for all output steps. However, because behaviour change was not measured this approach dropped out.

Instead, all values were standardised by converting them into proportions. Attitudes scores had to be modified slightly by converting the 1-5 Likert scale into a 0-4 scale before transforming summed values into proportions. After this, specific weights were allocated for each output step. Following the sequence described by McGuire, exposure is found as the first output step in the persuasiveness sequence and, as discussed by McGuire, would have the weakest effect on behaviour change. Thus, a weight of 0.25 was given to the exposure output step. Attention (two methods) and

comprehension-lexical meaning were measured together but because they still have a stronger effect than exposure on the communication sequence, they were given a weight of 0.35. Comprehension-psychological meaning was also used as a measure of comprehension, but because it is not such a simple measure as comprehension-lexical meaning, a score of 0.40 was given. Following McGuire's sequence, attitude is the output step that follows and therefore was given a weight of 0.5. It was difficult to decide which step followed attitudes, and while McGuire has put skills acquisition as the next step, a personal communication with Elder lead to the choice of skills acquisition as the last and most important output step because if mothers did not acquire the skills to perform what they were taught, then all the previous output steps would be irrelevant. In the end, it is more important to measure what a person can actually do, than what she says she can do. Hence, recall was given a weight of 0.55 and skills, a weight of 0.60.

For each of the intervention groups, weights were multiplied by each output step value (or % value). An overall score was generated for each intervention group. A t-test was used to compare mean scores. For the exposure output step which measured intervention II as GC and IC, an average of both groups was calculated before giving the weights.

4.8.6 Software packages

Double data entry for each of the output steps (except for exposure where Excel was instead used) was carried out by the main researcher and was undertaken using CSPro version 3.0^[178]. Raw data files were converted from CSPro into SPSS 14.0 which was used for data management purposes and all statistical analyses.

CHAPTER 5

RESULTS

This chapter presents the study results for each of the output steps and compares them between the intervention groups. This chapter also presents any associations between output steps and exposure variables, and any associations among output steps.

5.1 Exposure

Even though the term exposure, as defined by William McGuire has been the term used in previous chapters, in this section and throughout the rest of the thesis it will be called attendance since exposure in the current research study has been measured as attendance to classes. Attendance rates were calculated as the proportion of total classes attended.

5.1.1 Characteristics of study subjects

Whilst all the other output step measurements were conducted in a sub-sample of the TSNI study population, attendance measurements were conducted in all of the TSNI intervention I and II study population.

P-P plots (not shown) for the cumulative proportion group class (GC) attendance and individual class (IC) attendance against the cumulative proportion of a normal distribution have shown points clustering around the normal distribution straight line for GC attendance and a large deviation from normality for IC attendance.

Table 5.1 describes those characteristics that could have a potential influence on the mothers' attendance levels. As shown in the table, characteristics were very similar and not significantly different, as measured by an Independent-Samples T-test, between the two intervention groups. Not surprisingly, education levels of the mothers were extremely low, with the majority of the mothers in both intervention groups having only completed the first grade in the education system. As shown in the table, around 75% of the mothers in both intervention groups perceived their children (TSNI's reference children) to be in good health and growing well, in contrast baseline results showed

that mean weight for age and height for age for both interventions was around -2 Z-scores .

Table 5.1 Characteristics of TSNI study population by intervention group (n=560)

CHARACTERISTIC ¹	MEAN VALUE (SD)		P-VALUES
	INTERVENTION I	INTERVENTION II	
Age of mothers (age in years)	27.62 (7.31) (n=274)	27.28 (6.99) (n=286)	0.574
Education level of mothers (years of schooling)	1.02 (1.59) (n=274)	1.24 (1.85) (n=286)	0.139
Education level of father of reference child (years of schooling)	3.03 (2.48) (n=237)	3.07 (2.42) (n=242)	0.856
Household size	5.18 (1.59) (n=274)	5.10 (1.69) (n=286)	0.614
Income per capita (thousand MZM) ²	584.31 (368.41) (n=246)	600.06 (487.95) (n=252)	0.684
% of mothers that perceived their child to be growing well and to be in good health	76 (n=254)	73 (n=258)	0.730

¹ All characteristics were measured at baseline (May 2003)

² MZM= Meticalis – Mozambique's currency (1USD = approx. 24.000 MZM)

5.1.2 Attendance rates to group and individual classes

Table 5.2 presents the descriptive statistics for the attendance to either group classes (GC) or individual classes (IC) by intervention group (i.e. mean number of classes attended). Because total number of group classes provided by TSNI was 12 and individual classes was 6, the table also presents the descriptive statistics for the proportion (in percentage) of attendance to each approach-intervention group (i.e. GC-intervention I, GC intervention II and IC). As shown in the table, mean proportion of attendance (%) to IC was the highest.

Table 5.2 Descriptive statistics for attendance by approach-intervention group

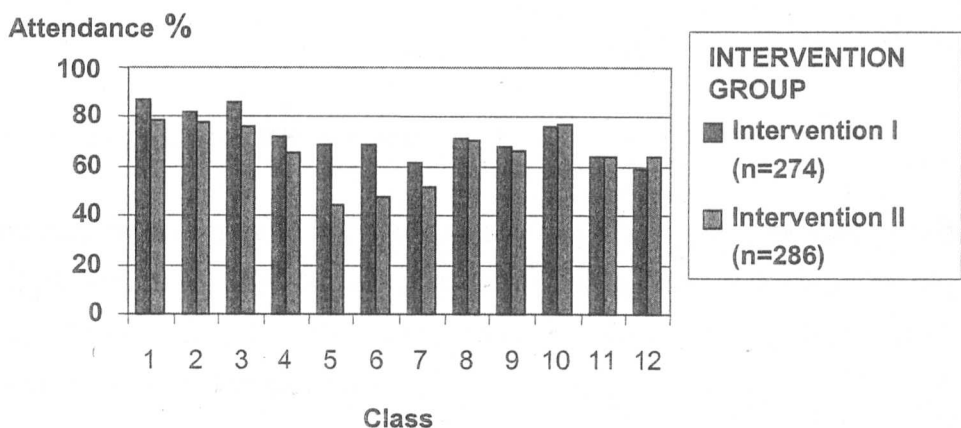
Approach- Intervention group	Mean	N	SD	Minimum	Maximum
GC intervention I	8.65	274	3.229	0	12
GC intervention I (% attendance)	72.08 %	274	26.908	0	100
GC intervention II	7.84	286	3.401	0	12
GC intervention II (% attendance)	65.33%	286	28.342	0	100
IC	5.38	286	1.645	0	6
IC (% attendance)	89.69%	286	27.421	0	100

5.1.2.1 Group class attendance

Overall 12 group classes were provided by the TSNI project and attendance at each of these classes was recorded for all TSNI study mothers (i.e. both from intervention I and II).

Figure 5.1 shows attendance level for each specific class in chronological order (i.e. overall attendance rate on the first topic, then overall attendance level on the second topic, and so on).

Figure 5.1. Attendance level at each class (topic) by intervention group

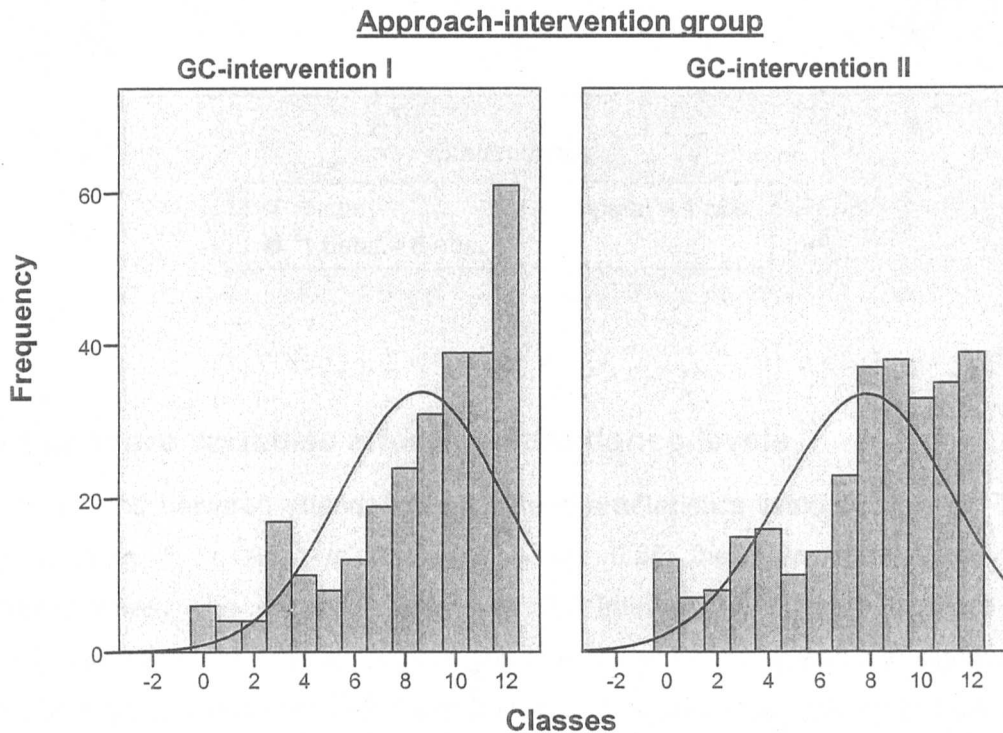


As shown in figure 5.1, attendance levels varied over the 12 classes, started by being higher and lowered at the end of the program. This difference between the attendance

to the first class and attendance to the last class was especially pronounced for intervention I (Paired-Samples T-Test, mean diff 0.14; CI: 0.004 – 0.273, $p=0.043$).

Figure 5.2 on the other hand, shows overall number of classes attended by mothers for each GC-intervention group. As shown in the figure, a large proportion of mothers attended most of the classes. This was especially the case for intervention I mothers.

Figure 5.2 Total number of group classes attended by approach-intervention group

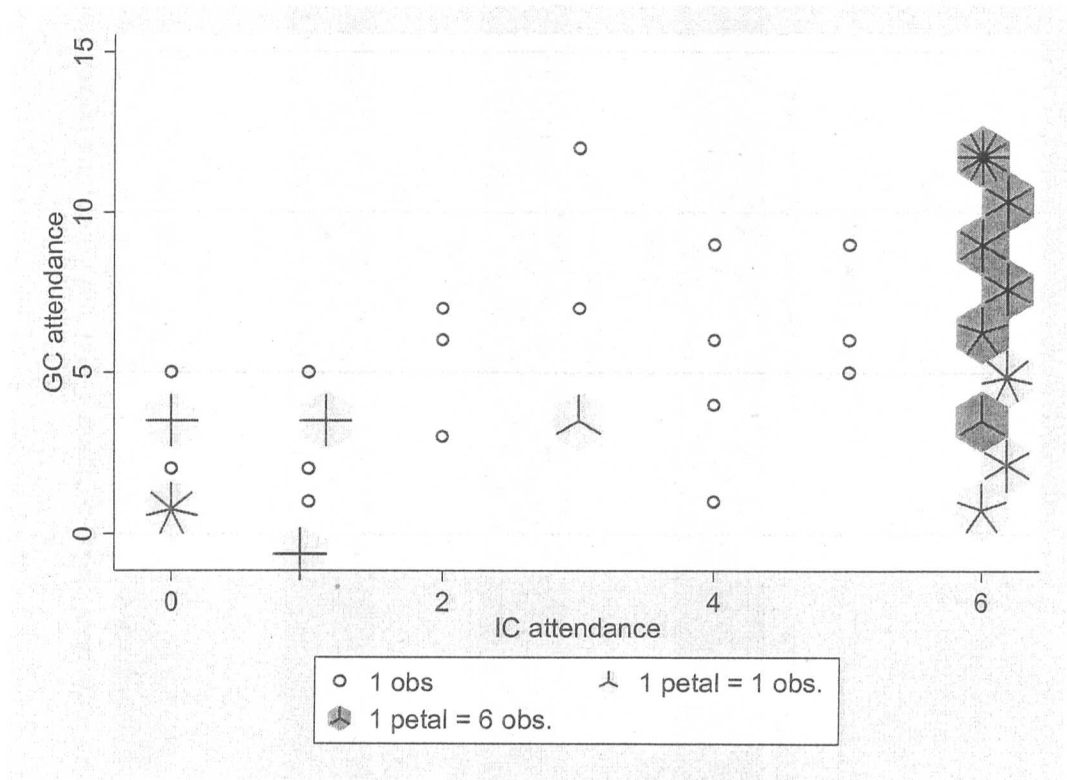


5.1.2.2 Individual class attendance

Regarding attendance rates to individual classes, in total 6 classes were provided by the TSNI project and attendance to each of these was recorded for all TSNI intervention II study subjects. Attendance rates were very high with 85.7% of mothers attending all the 6 individual classes.

To see if there is an association between attending individual classes and also group classes, attendance to both was correlated. Figure 5.3 depicts this association. There is a strong correlation (Pearson's test, $r= 0.573$, $p<0.001$) with mothers attending fewer individual classes also attending fewer group classes.

Figure 5.3 Correlation between Intervention II attendance to group and individual classes



5.1.3 Exposure variables affecting attendance levels

The correlations between attendance and the characteristics listed in table 5.1 were explored (table 5.3). All r-values were below 0.30 therefore correlations were considered weak. The exception was with GC intervention I where mothers who believed their reference children to be in good health and growing well had significantly higher attendance levels than those that believed their reference children to have poor health and growth.

Since table 5.3 shows multiple associations, there is a possibility of finding an association because of chance (type 1 error) but there is a also a possibility of making type II error – not finding an association when there is one.

Table 5.3 Associations between Socio-Demographic variables and attendance to group and individual classes (using T-test and Pearson's correlation)

EXPOSURE	GC ATTENDANCE INTERVENTION I	GC ATTENDANCE INTERVENTION II	IC ATTENDANCE INTERVENTION II
Age of Mother	Pearson's r= .030	Pearson's r= .101 *	Pearson's r= .000
Education level of mother (years of schooling)	(n= 274) Pearson's r= -.134 ** CI: -.226 – 1.346	(n=286) Pearson's r= -.095 CI: -.184 – 1.443	(n=286) Pearson's r= .051 CI: -.594 -- .195
Mother has formal education? (Yes/No)	(n=274) Pearson's r= -.074	(n=286) Pearson's r= -.198 ***	(n=286) Pearson's r= -.139 **
Education level of father (years of schooling)	(n=237) Pearson's r=.042	(n=242) Pearson's r= .098 *	(n=242) Pearson's r= -.001
Household size	(n=274) Pearson's r= -.102	(n=286) Pearson's r= -.105 *	(n=286) Pearson's r=.025
Income per capita	(n=246) CI of difference: -2.272 -- -.208 **	(n=252) CI of difference : -.895 – 1.223	(n=252) CI of difference: -.455 -- .544
Mother's perception of child's health (growing well or poor growth)	(n=254) *** p< 0.01 **** p< 0.001	(n=258)	(n=258)
* p<0.10 ** p< 0.05			

5.1.4 Comparing attendance rates between approach-intervention groups

Table 5.4 Analysis of Variance of attendance rates

Approach-intervention group	Mean % (SD)	Mean Difference (95% CI)	Sig.
IC	89.69 (27.42)	-17.61 (-23.20, -12.01)	0.000
GC Intervention I	72.08 (26.91)		
GC intervention II	65.33 (28.34)	6.75 (1.16, 12.35)	0.012
GC Intervention I	72.08 (26.91)		

The mean difference is significant at the .05 level. Includes Bonferroni correction

Table 5.4 gives a detailed description of the overall differences in attendance rates across the 3 approach-intervention groups using an ANOVA test. It also shows a break down of results for the comparison between each of the approach-intervention groups. Attendance rates for IC are significantly higher than attendance rates in both GC intervention I and II. Attendance rates for GC intervention I were also significantly higher than attendance rates for GC intervention II.

When the comparison was between intervention groups, a T-test showed that proportion attendance to intervention II was significantly higher (mean diff. 5.43, CI: -9.72 -- -1.13; $p=0.013$) than that to intervention I.

5.2 Attention

Attention was measured using 10 direct observations during class and by using 2 recall questions asked at the end of the class. Unfortunately, as the identity of mothers whose attention was measured in group classes is not known, it is not possible to compare demographic characteristics between the two intervention groups.

P-P plots (not shown) for the cumulative proportion of GC intervention I, GC intervention II and IC attention levels (from observations) against the cumulative proportion of a normal distribution have shown points clustering around the normal distribution straight line for both GC attention observation scores but IC attention observation scores deviate relatively more from the normal distribution straight line.

Other exposure factors' characteristics are depicted in table 5.5. Measurements for these exposure factors were taken at the same time as attention and comprehension-lexical meaning measurements.

Table 5.5. Characteristics (mean and SD) of variables (i.e. exposure factors) measured during attention measurements

Approach-intervention groups		Number of women standing up	Number of children crying	Class size (number of adults)	Duration of class (in min)
GC intervention I	Mean	1.76	2.38	13.43	35.00
	N (n)	5 (42)	4 (32)	5 (42)	3 (22)
	Std. Deviation	.983	2.03	1.755	21.97
GC intervention II	Mean	3.19	3.80	14.53	49.32
	N	10 (57)	10 (51)	10 (57)	5 (34)
	Std. Deviation	2.158	3.25	5.145	10.55
IC	Mean	Not relevant	3.30	Not relevant	29.30
	n		91		60
	Std. Deviation		2.5		11.82

N = number of group classes

Mean class duration was significantly higher for GC intervention II when compared with IC (mean diff. 20, $p < 0.001$) and GC intervention I (mean diff. 14.32, $p = 0.001$), however as can be seen from the above table, class duration was only measured for a few classes, reducing the sample size. Only 22 out of 42 mothers were measured in GC intervention I, 34 out of 57 in GC intervention II and 60 out of 91 in IC.

Mean number of women standing up and mean number of children crying was also significantly higher for GC intervention II when compared to GC intervention I ($p < 0.001$ and $p = 0.029$, respectively). It was not possible to compare the mean number of children crying in the GC approach with IC as this variable was measured in different ways in the different approaches.

This analysis thus provides two exposures, namely class duration and disturbance levels (i.e. women standing up and children crying) which could confound any associations between approach-intervention groups and attention scores.

In order to measure the test-retest reliability of both attention methods, attention levels in IC was measured twice for 10 mothers. Table 5.6 show attention levels for method 1 and method 2, on two different occasions. Pearson's correlation was used to measure test-retest reliability for both methods. A minimum coefficient of 0.70 is recommended

in research.[116][117] With the observation method the test-retest coefficient was 0.423 ($p=0.224$) which is considered weak, while with the recall question method the test-retest coefficient was 0.655 ($p=0.04$) which can be considered strong.

Table 5.6 Attention levels measured with both methods in 10 cases on two occasions

Cases	Attention observation (method 1)		Attention question (method 2)	
	First measurement	Second measurement	First measurement	Second measurement
1	10	7	2	2
2	10	9	1	1
3	6	8	1	2
4	10	10	2	2
5	9	8	1	1
6	8	10	1	1
7	10	10	2	2
8	6	6	1	2
9	7	9	2	2
10	9	8	2	2
Total	n	10	10	10

5.2.1 Exposure factors affecting attention

Exposure factors were associated with the two methods of attention measurement (Table 5.7). These exposure variables are those collected at the same time as the attention measurements and also socio demographic characteristics of IC mothers.

There was a negative medium to strong association between duration of group classes and observed attention (i.e. method 1) scores in GC intervention I and II. In contrast, there was a positive association between duration of group classes and attention recall (i.e. method 2) scores in GC intervention II, where scores were significantly higher for a higher duration of class. Duration of class did not seem to have an effect on attention scores from either method in IC.

It was observed that 64.3% of those mothers carrying a child were disturbed at least once by that child. Number of disturbances ranged from one to 10 times in one class/session, and the mean for both approaches was approximately 3. However, as shown in the table there was no association between child disturbance in group classes and attention scores measured with both methods. In fact, in IC, higher child disturbances was associated with a significantly higher recall of attention questions.

The frequency of mothers standing-up only had a weak association with GC intervention II attention scores measured with method 1. Group classes ranged from having 6 to 23 participants. However, group size did not seem to have any association with attention scores measured with either method for both GC intervention groups. Moreover, none of the socio-demographic variables seemed to have any association with attention levels measured with either method in IC.

As shown in the table there are a lot of missing values for those variables measured along with attention which might have affected results. It becomes difficult to interpret results with a smaller sample size, at it might not be representative.

Table 5.7 Association between exposure variables and attention levels (using Pearson's correlation test (*r*), Independent T-test (CI of difference of means), and chi-square test)

Variables	Attention in GC intervention I		Attention in GC intervention II		Attention in IC	
	Method 1	Method 2	Method 1	Method 2	Method 1	Method 2
Duration of class (minutes)	Pearson <i>r</i> =-.622 *** (n=22)	CI: -15.202 -- 34.916 (n=21)	Pearson <i>r</i> =-.396 ** (n=28)	CI: -17.752 -- -3.616 *** (n=27)	Pearson's <i>r</i> = .232 * (n=54)	CI: -5.788 -- 10.329 (n=54)
Frequency of crying of children	Pearson's <i>r</i> = -.100 (n= 32)	CI: -2.036 -- .693 (n=31)	Pearson's <i>r</i> = -.173 (n= 51)	CI: -3.111 -- .127 * (n=51)	Pearson's <i>r</i> = -.036 (n=64)	CI: -2.373-- .040 ** (n=64)
Frequency of standing ups	Pearson's <i>r</i> = -.214 (n= 42)	CI: -.079 -- .693 * (n=41)	Pearson's <i>r</i> = -.282 ** (n= 57)	CI: -.516 -- 1.898 (n=56)	NM	NM
Size of group class	Pearson's <i>r</i> = -.092 (n= 42)	CI: -1.365 -- .790 (n=41)	Pearson's <i>r</i> = .086 (n= 57)	CI: -.250 -- 6.268 * (n=56)	NA	NA
Age of Mother (in years)					Pearson's <i>r</i> = .149 (n=79)	CI: -3.175--3.581 (n=79)
Education level of mother (years of schooling)					Pearson's <i>r</i> = -.059 (n=79)	CI: -1.764--.447 (n=79)
Mother has formal education? (yes/No)					CI: -920 --1.112 (n=78)	Chi-square p=0.409 (n=79)
Education level of father (years of schooling)					Pearson's <i>r</i> = -.017 (n=68)	CI: -1.606--1.816 (n=68)
Income per capita					Pearson's <i>r</i> = .007 (n=74)	CI: -256.874 -- 225.606 (n=74)

Method 1 = Attention observations and Method 2 = recall-type attention questions
 the ID of mothers in GC intervention I and II was not known
 * *p*<0.10 ** *p*< 0.05 *** *p*< 0.01 **** *p*< 0.001
 NA= not measured Shaded boxes – not possible to associate as

5.2.2 Comparing observed attention scores between intervention groups and between approach-intervention groups

Overall 10 direct attention observations of eye contacts between mother and health promoter were carried out per session at both the group classes and the individual classes.

Table 5.8 presents mean attention scores by approach-intervention group. The mean IC attention score was approximately 11% higher than that for GC intervention I. However, the mean GC intervention I attention score was slightly higher than that for GC intervention II. This difference in scores can also be viewed in figure 5.4.

Table 5.8. Descriptive statistics for attention observation by approach-intervention group

Approach-intervention group	N	Mean (SD)	95% Confidence Interval for Mean		Min	Max
			Upper Bound	Lower Bound		
GC-Intervention I	42	7.52 (2.18)	6.85	8.20	2	10
IC	91	8.36 (2.11)	7.92	8.80	0	10
GC--Intervention II	57	7.46 (1.91)	6.95	7.96	3	10
Total	190	7.91 (2.10)	7.60	8.21	0	10

Figure 5.4. Observed attention scores by approach-intervention group

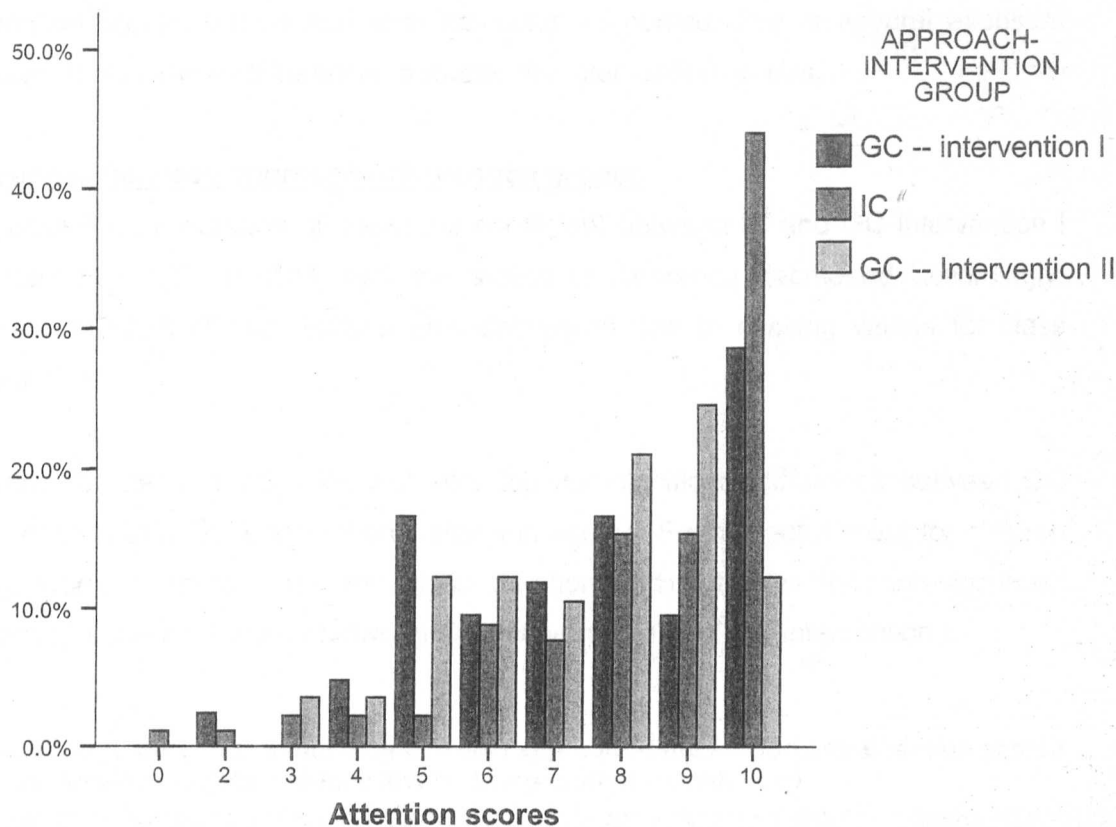


Table 5.9 gives a detailed description of the overall differences in observed attention scores across the 3 approach-intervention groups using an ANOVA test. It also shows a break down of results for the comparison between each of the approach-intervention groups. There was no difference between IC and GC intervention I and no difference between GC intervention II and GC intervention I. The only significant result was between IC and GC intervention II, where attention scores were higher in IC.

Table 5.9 Analysis of Variance for attention scores

Approach-intervention group	Mean (SD)	Mean Difference (95% CI)	Sig.
IC	8.36 (2.11)		
GC Intervention I	7.52 (2.18)	-0.839 (-1.77, 0.09)	0.093
GC intervention II	7.46 (1.91)		
GC Intervention I	7.52 (2.18)	0.068 (-0.95, 1.08)	1.000
IC	8.36 (2.11)		
GC Intervention II	7.46 (1.91)	-0.906 (-1.751, -0.06)	0.031

The mean difference is significant at the .05 level. Bonferroni correction included.

Table 5.10 shows differences in attention observation scores between approach-intervention groups before and after adjusting for confounding of several exposure variables. It also shows differences between the intervention groups.

Differences between approach-intervention groups

After adjusting for duration of class, the coefficient between IC and GC intervention I increased from .82 to 1.03, and the statistical difference decreased accordingly. However, number of observations also decreased due to missing values for class duration.

Adjusting for class duration did not alter the non-significant difference between GC intervention II and GC intervention I attention scores. Further adjustment for children crying, women standing up and class duration, did not alter the non-significant difference in attention scores between GC intervention II and GC intervention I.

Table 5.10 Adjusting the association between approach-intervention and attention scores for class duration and disturbance levels (using multiple regression)

Model	Approach-intervention group	Coefficient	P> z
Crude (not adjusting for confounding)	GC intervention I	1	
	IC	.82	0.031
	GC intervention II	-.07	0.872
Number of obs=190	INTERVENTION II	.49	0.184
Adjusting for class duration	GC intervention I	1	
	IC	1.03	0.062
	GC intervention II	.064	0.919
Number of obs= 116	INTERVENTION II	.70	0.178
Adjusting for children crying and women standing up	GC intervention I	1	
	GC intervention II	.75	0.134
Number of obs=83			
Adjusting for class duration, women standing up and children crying	GC intervention I	1	
	GC intervention II	.62	0.330
Number of obs=44			

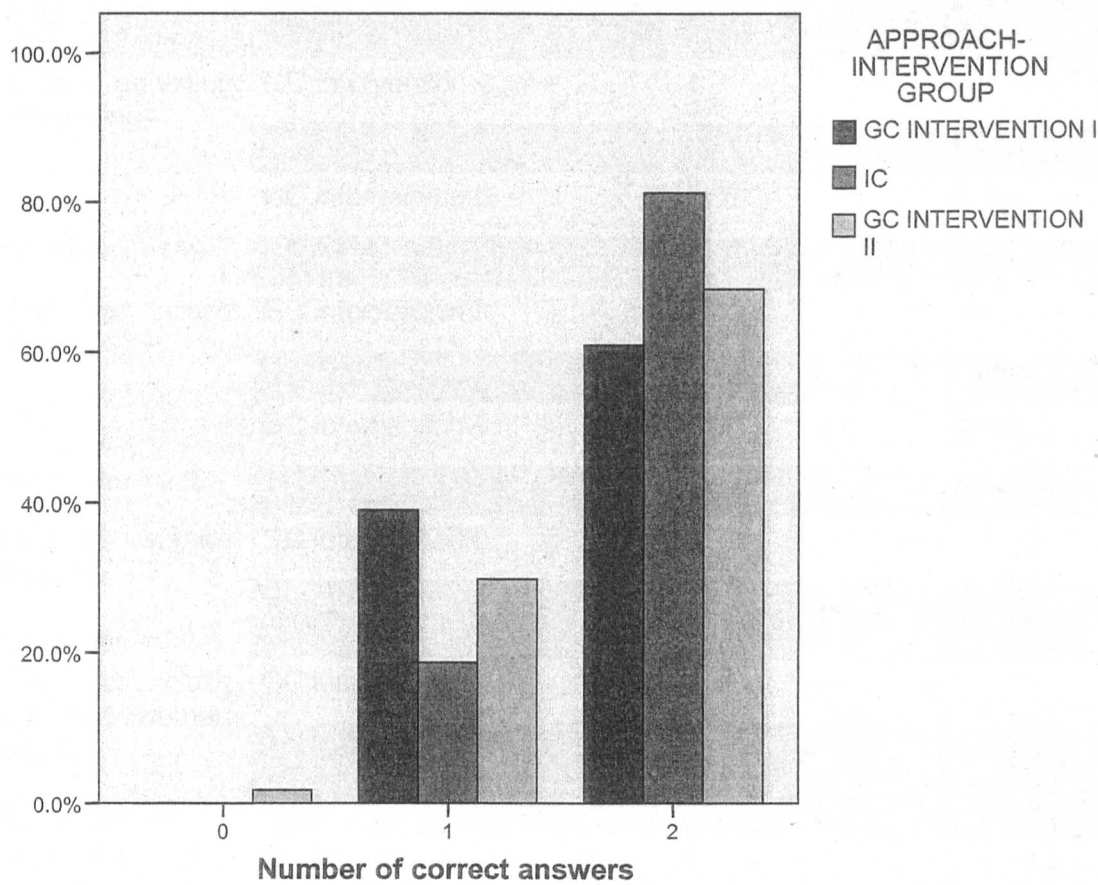
Differences between intervention groups

After adjusting for duration of class, the coefficient between Intervention II and intervention I increased slightly, and the statistical difference remained not significant.

5.2.3 Comparing results for attention-recall between intervention groups and between approach-intervention groups

Attention was also measured by asking 2 simple recall-type questions immediately after the class. The first question was a bogus recall-type question of something that was not mentioned in the class, while the second question was a recall-type question of something that was mentioned in the class. Figure 5.5 shows that there was a higher proportion of correct answers to two questions (hence, higher recall) in IC compared to the other approach-intervention groups.

Figure 5.5 Attention-recall by approach-intervention group



Only 1 mother in GC intervention II answered both recall questions incorrectly. To be able to make this variable into a binary one, this result was joined with the group of mothers that answered correctly 1 question.

Pearson's Chi-square was used to compare approach-intervention groups. There was a 20% significantly higher ($p=0.012$) proportion of correct answers to two recall questions in IC mothers compared to GC intervention I mothers. Even though there was a 10% higher proportion of correct answers to the two recall questions in GC intervention II in comparison to GC intervention I, this difference was not statistically significant ($p=0.455$).

Table 5.11 shows the associations between attention-recall scores in IC with GC-intervention I and between GC-intervention II with GC-intervention I before and after adjusting for confounding for several exposure variables. It also shows the differences between the intervention groups before and after adjusting for class duration.

Table 5.11 Adjusting the association between approach-intervention and attention-recall for class duration and disturbance levels (using logistic regression)

Model	Approach-intervention group	Odds Ratio (OR)	P> z
Crude (not adjusting for confounding)	GC intervention I	1	
	IC	2.78	0.014
	GC intervention II	1.38	0.446
Number of obs = 189	INTERVENTION II	2.06	0.052
Adjusting for class duration	GC intervention I	1	
	IC	2.95	0.069
	GC intervention II	.95	0.940
Number of obs = 115	INTERVENTION II	1.86	0.240
Adjusting for women standing up	GC intervention I	1	
	GC intervention II	2.05	0.137
Number of obs = 98			
Adjusting for class duration and women standing up	GC intervention I	1	
	GC intervention II	2.56	0.282
Number of obs = 49			
Adjusting for class duration, women standing up and children crying	GC intervention I	1	
	GC intervention II	1.79	0.809
Number of obs = 43			

Differences between approach- intervention groups

After adjusting for class duration, the OR between IC and GC intervention II increased only slightly, but the statistical difference between the two approach-intervention groups decreased. This decrease could have been caused by a decrease in the number of observations.

After adjusting for class duration, the OR between GC intervention II and GC intervention I decreased, and the statistical difference remained not significant.

After adjusting for class duration and disturbance levels (i.e. women standing up and children crying) the OR between GC intervention II and GC intervention I increased, but the statistical difference remained not significant.

Both class duration and women standing up were associated with increased odds of answering the attention-recall questions correctly. The OR for IC increased when class duration was adjusted for, suggesting that as class duration increased so did IC mothers odds of getting the attention-recall questions correctly. Similarly, the OR for GC-intervention II increased when women standing up was adjusted for, hence, as more women stood up per class so did GC-intervention II mothers odds of getting the attention-recall questions correctly.

Differences between the intervention groups

After adjusting for class duration the OR between intervention II and intervention I decreased and the magnitude of the difference also decreased and became statistically non-significant. This suggests that duration of class was confounding the relationship and that after adjustment no difference was observed between intervention I and II. However, in these analyses the sample size was considerably decreased.

5.3 Understanding/Comprehension

5.3.1 Comprehension- lexical meaning

Comprehension-lexical meaning was assessed via response to one question administered to the mother(s) right after the group and individual classes.

5.3.1.1 Exposure variables affecting comprehension-lexical meaning

Exposure variables depicted in table 5.5 are also relevant for this output step.

The association between comprehension-lexical meaning and the exposure variables were examined (table 5.12). The only statistically significant association was with child disturbance in GC intervention II, whereby child disturbance was significantly higher for the group that missed the understanding question.

Table 5.12 Association between understanding and exposure variables for each approach-intervention group (using Independent T-test (CI for the difference in the mean) and Chi-square)

Variables	Understanding in GC intervention I	Understanding in GC intervention II	Understanding in IC
Duration of class (minutes)	CI: -15.48, 25.76 (n=21)	CI: - 6.59, 8.13 (n=28)	CI: -1.84, 11.634 (n=54)
Frequency of children crying	CI: -1.58, 1.50 (n=31)	CI: 0.76, 4.41 *** (n=51)	CI: -1.23, 1.33 (n=64)
Frequency of women standing up	CI: -0.64, 0.63 (n=41)	CI: -0.42, 2.14 (n=55)	NM
Size of group class	CI: -1.11, 1.16 (n=41)	CI: -3.52, 2.36 (n=57)	NA
Age of mother (years)	Shaded area-- not possible to associate as the ID of mothers in GC intervention I and II was not known		CI: -2.14, 4.75 (n=79)
Mothers education level (years of schooling)			CI: -1.61, 0.16 (n=79)
Mother has formal education? (yes/No)			Chi-square, p=0.131 (n=79)
Education level of father (years of schooling)			CI: -1.43, 1.21 (n=68)
Income per capita			CI: -234.15, 155.17 (n=74)

NM = not measured NA= not applicable Shaded area-- not possible to associate as the ID of mothers in GC intervention I and II was not known

* p<0.10 ** p< 0.05 *** p< 0.01 **** p< 0.001

5.3.1.2 Comparing comprehension- lexical meaning between approach-intervention groups

Overall there were 3 types of comprehension questions which should have been administered to each intervention group and in equal proportions. However, because the GC approach was broken down into GG intervention I and GC intervention II, the questions were also broken down. The results is that question 1 was not administered to GC intervention II and all approach-intervention groups received different proportions of the 3 questions, for example, almost 50% of the intervention I mothers received question 1, while other 40% received question 2 and only 13% received question 3.

Regarding the difficulty of the questions, mothers in all approach-intervention groups seemed to have more difficulty in answering question 3 (not shown), as approximately 50% of answers for this question were incorrect, while in comparison only 25% for question 1 and approximately 21% for question 2 were incorrect.

Figure 5.6 Proportion of question type administered to mothers in each of the Approach- intervention groups.

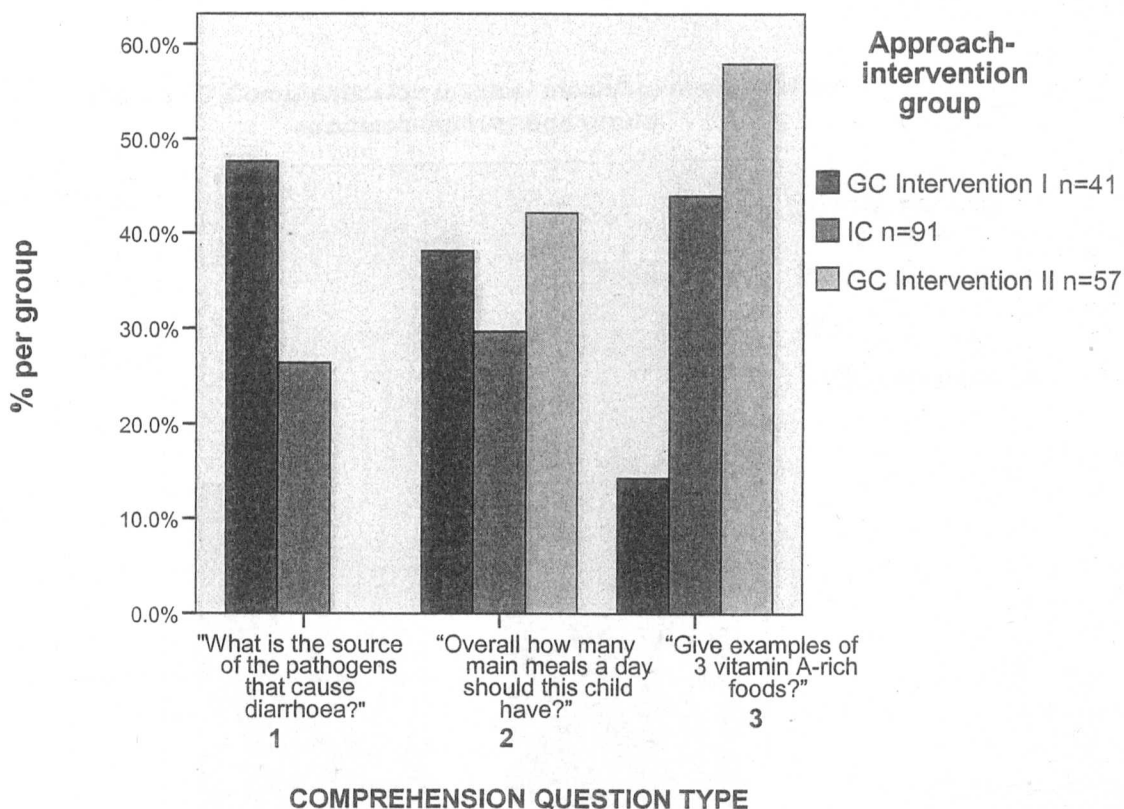


Figure 5.7 compares the overall (i.e. 3 questions pooled) proportion of correct answers by approach-intervention groups. There is a higher proportion of correct answers for IC.

Pearson's Chi-square was used to compare overall percent of correct answers between approach-intervention groups. There were 15.3% more correct answers for IC in comparison to GC intervention I but this difference was only of marginal significance (Pearson's Chi-square test, $p=0.084$).

There was no significant difference between GC intervention I and GC intervention II (Pearson's Chi-square test, $p=0.377$), and no significant difference between IC and GC intervention II (Pearson's Chi-square test, $p=0.405$).

On the other hand, because the question type and the number of times they were administered differed significantly (table 5.8) between the approach-intervention groups and the answers to the questions also differed significantly (chi-square, $p=0.001$) between the question type, question type may be confounding the association between approach-intervention groups and answers to the comprehension question and it might be that the lack of association between intervention groups and answer to comprehension questions arises from this confounder.

Figure 5.7 Comprehension (Lexical meaning) responses by approach-intervention group

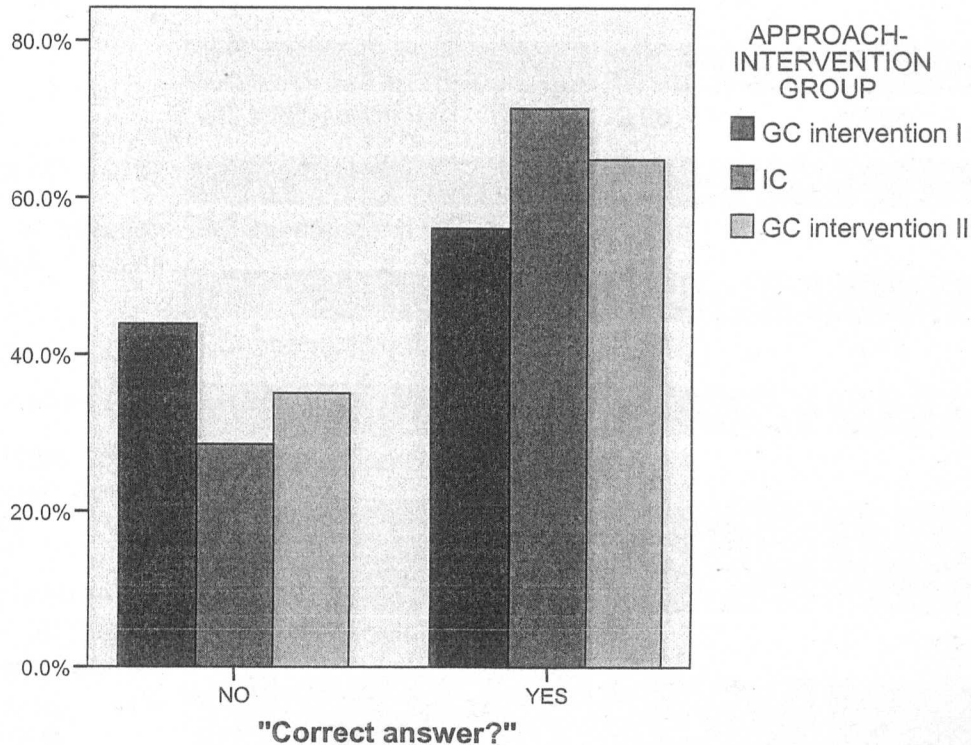


Table 5.13 looks at the association between answers to the comprehension question in IC with GC-intervention I and between GC-intervention II with GC-intervention I before and after adjusting for confounding of several exposure variables, including question type. It also looks at the difference between the intervention groups before and after adjusting for several exposure variables.

Table 5.13 Adjusting the association between approach-intervention groups and comprehension scores for class duration and disturbance levels (using logistic regression)

Model	Approach-intervention group	Odds Ratio (OR)	P> z
Crude (not adjusting for confounding)	GC intervention I	1	
	IC	1.95	0.086
	GC intervention II	1.45	0.378
	INTERVENTION II	1.73	0.127
(Number of obs =189)	GC intervention I	1	
	IC	3.62	0.005
	GC intervention II	3.57	0.011
	INTERVENTION II	3.60	0.003
Adjusting for question type	GC intervention I	1	
	IC	2.35	0.110
	GC intervention II	2.96	0.088
	INTERVENTION II	2.53	0.068
(Number of obs =115)	GC intervention I	1	
	IC	2.44	0.099
	GC intervention II	3.20	0.074
	INTERVENTION II	2.65	0.058
Adjusting for question type and class duration	GC intervention I	1	
	IC	6.53	0.029
	GC intervention II		
	INTERVENTION II		
(Number of obs =115)	GC intervention I	1	
	IC	9.13	0.020
	GC intervention II		
	INTERVENTION II		
Adjusting for class duration and women standing up	GC intervention I	1	
	IC		
	GC intervention II		
	INTERVENTION II		
(Number of obs =49)	GC intervention I	1	
	IC		
	GC intervention II		
	INTERVENTION II		
Adjusting for class duration and women standing up and children crying	GC intervention I	1	
	IC		
	GC intervention II		
	INTERVENTION II		
(Number of obs =43)	GC intervention I	1	
	IC		
	GC intervention II		
	INTERVENTION II		

Differences between approach-intervention groups

After adjusting for question type, the odds of getting two correct answers was 3.6 times higher for IC when compared to GC intervention I, and this was statistically significant. In addition, the odds of getting two correct answers was almost 3.6 times higher for GC intervention II when compared to GC intervention I and this was also statistically significant.

After adjusting for duration of class, the odds decreased. The odds of getting two correct answers was 2.3 times higher for IC when compared to GC intervention I, but this was not statistically significant. In addition, the odds of getting two questions correctly was almost 3 times higher for GC intervention I when compared to GC intervention I, but this was only of marginal significance.

Results suggest that both duration of class and question type were confounding the relationship between answers to understanding question and approach-intervention group since OR increased for both GC intervention II and IC when adjusted for these two confounders. Moreover, after adjustment for both confounders a stronger difference was observed between GC intervention II and GC intervention I. A similar effect was also seen when comparing IC with GC intervention I, although the statistical difference was lower when the adjustment was for class duration.

Class duration and disturbance levels (women standing up and children crying) seemed to have independent effects. After adjusting for these 3 confounders the OR between GC intervention II and GC intervention I increased from 1.45 to 9.13. This suggests that class duration and disturbance levels were confounding the relationship and that after adjusting for these, a stronger difference is observed between GC intervention II and GC intervention I.

Differences between intervention groups

Before adjusting for question type, there was no difference between the intervention groups. After adjusting for question type, intervention II proportion of correct answer to the comprehension question was significantly higher (OR: 3.60, $p=0.003$), than intervention I. After adjusting for class duration, intervention II proportion of correct answer to the comprehension question was marginally higher (OR: 2.53, $p=0.068$), than intervention I. After adjusting for both class duration and question type, intervention II proportion of correct answer to the comprehension question was significantly higher (OR: 2.65, $p=0.058$), than intervention I.

5.3.2 Comprehension- operational and psychological meaning

Comprehension was also measured as Comprehension- operational and – psychological meaning, and this was composed of questions that measured more elaborate and long term understanding. These were administered to the mothers after they had carried out the skills acquisition task.

P-P plots (not shown) for the cumulative proportion of intervention I and intervention II comprehension levels against the cumulative proportion of a normal distribution have shown points clustering around the normal distribution straight line for both intervention groups.

5.3.2.1 Exposure variables affecting comprehension- psychological and - operational meaning

The table below describes the means (SDs) for socio demographic characteristics and other exposures associated with block 3 (i.e. comprehension, skills acquisition and recall) output steps.

Apart from the father's educational level which was significantly higher for intervention I mothers, all other socio demographic variables did not differ significantly between the intervention groups.

Time interval (between recipe demonstration and the recipe task) was significantly higher for intervention I and recipe type (which is not shown in the table) also significantly differed (Chi-square, $p=0.055$) between the intervention groups.

Table 5.14 Descriptive statistics for exposure variables and comparison between intervention groups (mean difference, and confidence interval of the difference)

EXPOSURE VARIABLE	INTERVENTION I – MEAN (SD) n=40	INTERVENTION II – MEAN (SD) n=40	Differences between Intervention I & II
Age of Mother (age in years)	28.13 (7)	27 (6.4)	T-test Mean diff 1.13 CI: -2.07, 4.02
Education level of mother (years of schooling)	0.93 (1.46)	0.58 (1.15)	T-test Mean diff 0.35 CI: -0.23, 0.93
Education level of father (years of schooling)	3.08 (2.3)	1.91 (2)	T-test Mean diff 1.17** CI: 0.15, 2.20
Income per capita (thousands of MZM)	598.74 (378.3)	458.74 (288.2)	T-test Mean diff 140 CI: -35.03, 269.91
Time interval (days)	73.55 (20.78)	62.55 (26.12)	T-test Mean diff 11** CI: 0.48, 21.52
Cooking method	15 steaming: 25 boiling ¹	14 steaming: 26 boiling ¹	Chi-square, p=0.816
Number of ingredients	2.70 (0.939)	2.75 (0.981)	T-test Mean diff -0.05 CI: -0.48, 0.38

¹number of mothers that chose a cooking method *p<0.1 **p<0.05 ***p<0.01 ****p<0.001

Comprehension-psychological-operational meaning was then associated with the exposure variables depicted in the above table, the results of which are presented in table 5.15. There were no associations between comprehension levels and socio-demographic variables. Comprehension scores of intervention I mothers' were strongly associated with cooking method, whereby those mothers that chose the steaming method got significantly higher comprehension scores. Comprehension scores of intervention I mothers' also had medium association with type of recipe and number of ingredients used in the recipe.

Table 5.15. Association between comprehension-psychological meaning levels and exposure factors

EXPOSURE FACTORS	COMPREHENSION SCORES	
	INTERVENTION I (n=40)	INTERVENTION II (n=40)
Age of Mother (in years)	Pearson r= .159	Pearson r=.011
Education level of mother (years of schooling)	Pearson r= .294 *	Pearson r= -.115
Mother has formal education? (No/Yes)	T-test CI: -1.83, 37.57 *	T-test CI: -20.77, 15.81
Education level of father (years of schooling)	Pearson r= .174	Pearson r= .113
Time interval (days)	Pearson r= -.100	Pearson r= -.245
Cooking method (Steaming/boiling)	T-test CI: 7.19, 41.82 ***	T-test CI: -22.67, 12.54
Number of ingredients	Spearman's r= .435 ***	Spearman r= -.165
Recipe type	Spearman's r=0.404 ***	Spearman's r= -.128

*p<0.1 **p<0.05 ***p<0.01 ****p<0.001

5.3.2.2 Comparing comprehension- psychological and -operational meaning scores between the intervention groups

Table 5.16 Descriptive statistics for comprehension

	INTERVENTION GROUP	N	Mean	Std. Deviation	P-value
Comprehension score		40	54.99	29.39	0.376
	INTERVENTION I				
	INTERVENTION II	40	60.44	25.21	

Table 5.16 and figure 5.8 compare comprehension mean scores between the two intervention groups. Mean unadjusted scores were 10% higher for intervention II.

Figure 5.8 Comparing Comprehension (Psychological and Operational meaning) scores between intervention I and II

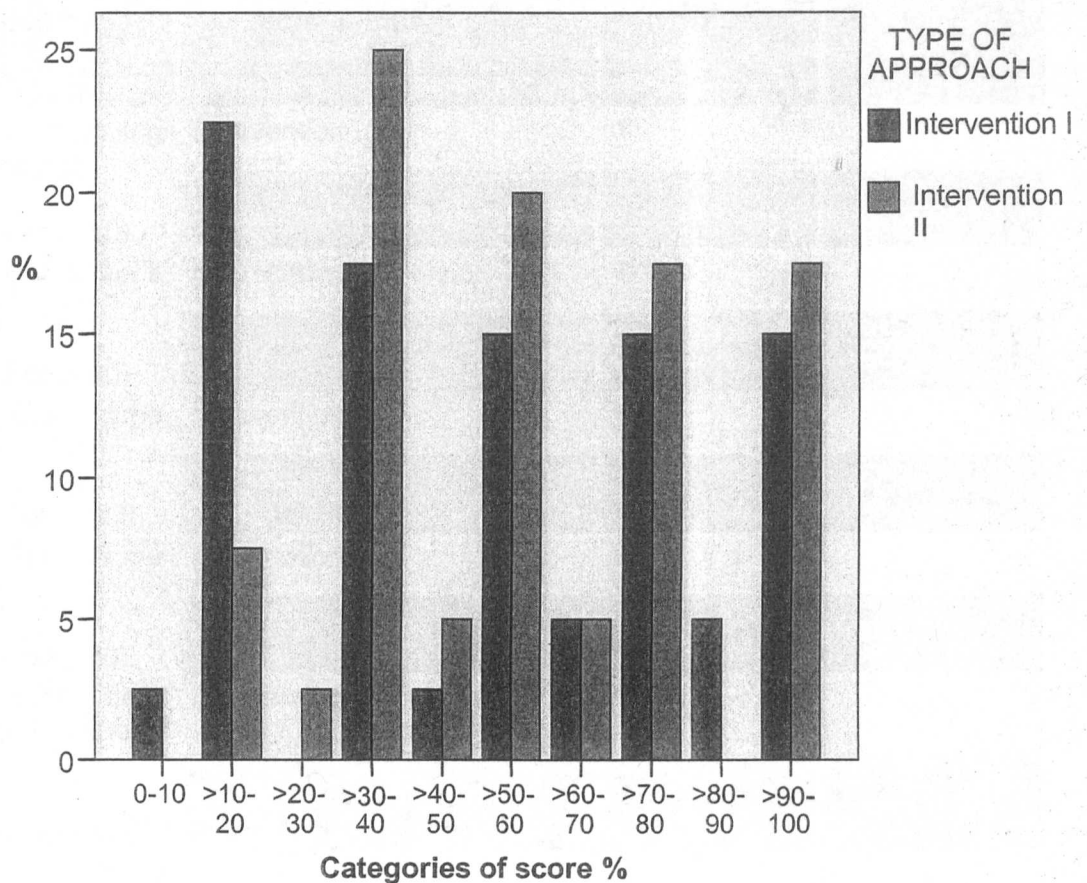


Table 5.17 compares comprehension scores between the intervention groups before and after adjusting for possible confounders.

After adjusting for fathers education, the coefficient for comprehension-psychological scores increased, putting scores in intervention II approximately 16% higher than intervention I, although the difference between the intervention groups remained non-significant.

After adjusting for time interval, the coefficient decreased and the statistical difference between the two intervention groups also decreased. Recipe type did not seem to be much of a confounder since the coefficient did not change much. After adjusting for both fathers education and time interval, the two potential confounders, the coefficient increased only slightly.

Table 5.17 Adjusting the association between intervention groups and comprehension scores for confounding (using regression)

Model	Intervention group	Coefficient	P> z
Crude (not adjusting for confounding)	Intervention I	1	
	Intervention II	5.44	0.376
(Number of obs=80)			
Adjusting for father's education	Intervention I	1	
	Intervention II	8.81	0.209
(Number of obs=80)			
Adjusting for time interval	Intervention I	1	
	Intervention II	3.25	0.604
(Number of obs=80)			
Adjusting for recipe type	Intervention I	1	
	Intervention II	5.68	0.353
(Number of obs=80)			
Adjusting for time interval and father's education	Intervention I	1	
	Intervention II	6.33	0.376
(Number of obs=80)			

5.4 Skills acquisition

Skills acquisition was composed of an assessment of 6 steps that the mothers had to follow when preparing the recipes. Each mother was taught 3 recipes in the initial demonstration, however, these 3 recipes were not necessarily the same for all the mothers (table 5.18).

Moreover, cooking methods varied according to the recipe but this was nonetheless similar between the intervention groups : 65% mothers in intervention I and 62.5% in intervention II chose those recipes that involved the boiling method.

Recipes chosen were also very similar between the intervention groups: 17.5% in intervention I and 15% in intervention II chose the easiest recipe (i.e. OFSP and coconut milk) and 32.5% in intervention I and 37.5% in intervention II chose the most difficult recipes (i.e. those with 4 ingredients). The preferred recipe for both

interventions was the one using toasted peanut, with 45% in intervention I and 47.5% in intervention II choosing this recipe.

Table 5.18 Recipe type and ingredients demonstrated per locality

Locality	ingredients -- Recipe 1		ingredients -- Recipe 2			ingredients -- Recipe 3			
	OFSP	Coconut milk	OFSP	Toasted peanuts		OFSP	beans	Coconut milk	Dark green leaves
Malei									
Posto Campo/luala									
Catale									
Mexixine				Coconut milk	egg				

All recipes with 3- 4 ingredients used the steaming method.

P-P plots (not shown) for the cumulative proportion of intervention I and intervention II skills acquisition levels against the cumulative proportion of a normal distribution have shown points clustering around the normal distribution straight line for intervention I and deviating relatively more for intervention II.

5.4.1 Exposure variables affecting skill acquisition scores

Table 5.19 presents the associations between skill acquisition scores and socio-demographic variables and other exposure variables. There was no association between skill acquisition scores and socio-demographic variables. The only significant association was a strong (negative) association between intervention II skills acquisition scores and time interval. Time interval did not seem to affect intervention I.

Table 5.19. Association between skills acquisition and exposure variables (using T-test (CI of the difference in means, Spearman's rho and Pearson's r)

EXPOSURE VARIABLES	SKILLS ACQUISITION SCORES*	
	INTERVENTION I (n=40)	INTERVENTION II (n=40)
Age of Mother (in years)	Pearson's r= -.021	Pearson's r= -.203
Education level of mother (schooling years)	Pearson's r= -.046	Pearson's r= .018
Mother has formal education? (Yes/No)	t-test CI: -9.68, 6.22	T-test CI: -0.77, 0.76
Education level of father (schooling years)	Pearson's r= .214	Pearson's r= .254
Income per capita (in MZM)	Pearson's r= -.070	Pearson's r= -.065
Cooking method used [vapour/boiling]	T-test CI: -1.11, 0.54	T-test CI: -1.11, 0.54
Recipe chosen [5 recipes]	Spearman's r= -.284 *	Spearman's r= .131
Number of ingredients used [2, 3 or 4]	Spearman's r= -.237	Spearman's r= .188
Time interval (in days)	Pearson's r= -.066	Pearson's r=-.576 ****

*p<0.1 **p<0.05 ***p<0.01 ****p<0.001

5.4.2 Comparing skill acquisition scores between intervention groups

Table 5.20 and figure 5.9 compare skills acquisition mean scores between the two intervention groups. Mean unadjusted scores were almost 25% higher for intervention II.

Table 5.20 Descriptive statistics for Skills acquisition scores

Intervention group		N	Mean	Std. Deviation	P-value
Skills scores	Intervention I	40	4.11	1.22	<0.001
	Intervention II	40	5.14	1.03	

Figure 5.9. Comparing skills acquisition scores between Intervention groups

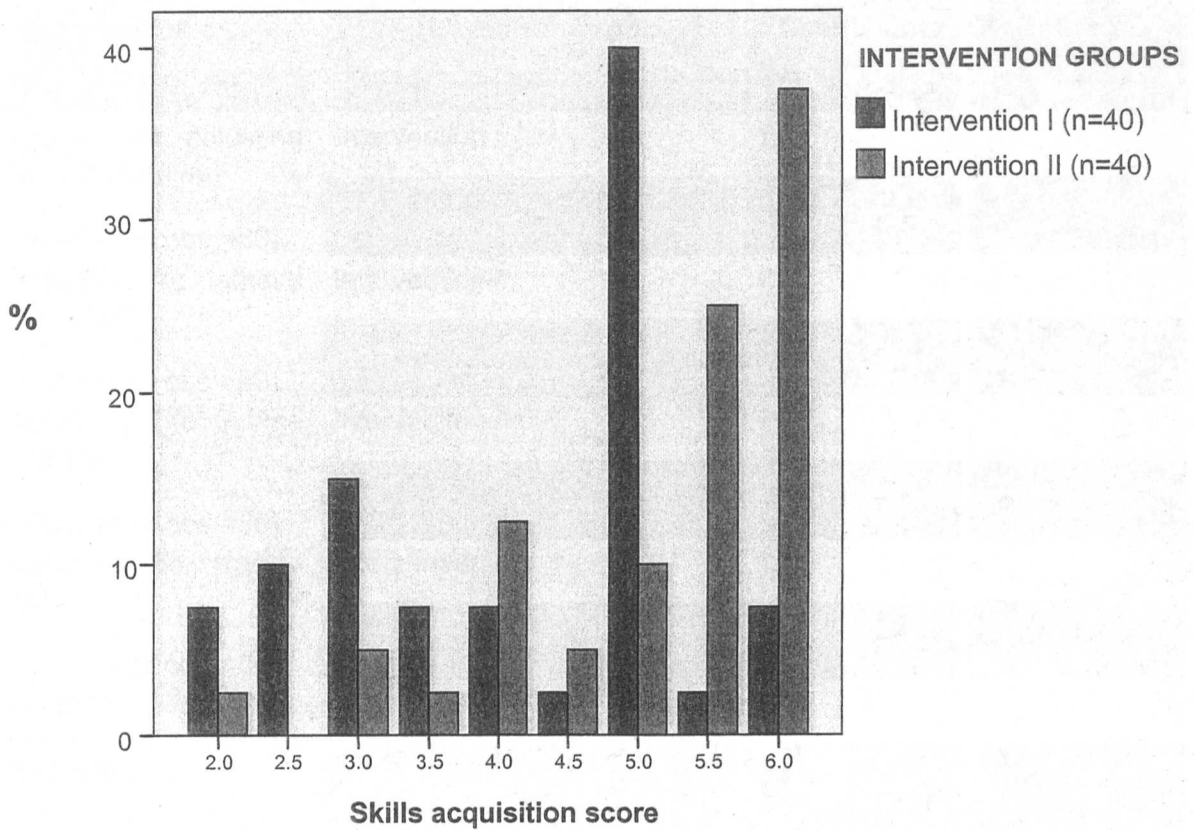


Table 5.21 looks at the association between recall scores in intervention I with intervention II before and after adjusting for confounding.

Even after adjusting for each confounder and combination of potential confounders, intervention II skills scores were significantly higher than intervention I.

Table 5.21 Adjusting the association between intervention group and skills acquisition scores for possible confounders (using regression)

Model	Intervention group	Coefficient	P> z
Crude (not adjusting for confounding) (Number of obs =80)	Intervention I	1	
	Intervention II	1.02	0.000
Adjusting for father's education (Number of obs =80)	Intervention I	1	
	Intervention II	1.17	0.000
Adjusting for time interval (Number of obs =80)	Intervention I	1	
	Intervention II	.85	0.001
Adjusting for recipe type (Number of obs =80)	Intervention I	1	
	Intervention II	1.01	0.000
Adjusting for time interval and father's education (Number of obs =80)	Intervention I	1	
	Intervention II	.96	0.001

5.5 Recall

Recall was assessed via response to 7-9 questions (depending on the number of ingredients in the chosen recipe).

P-P plots (not shown) for the cumulative proportion of intervention I and intervention II recall levels against the cumulative proportion of a normal distribution have shown points clustering around the normal distribution straight line for intervention I and deviating relatively more for intervention II.

5.5.1 Exposure variables affecting recall scores

Table 5.22 Association between recall and exposure variables (Using T-test ,CI for the difference in means, Spearman’s rho, and Pearson’s r)

EXPOSURE VARIABLES	RECALL SCORES	
	INTERVENTION I (n=40)	INTERVENTION II (n=40)
Age of Mother (in years)	Pearson’s r= .223	Pearson’s r= -.111
Education level of mother (schooling years)	Pearson’s r= .274 *	Pearson’s r= -.117
Mother has formal education? (Yes/No)	T-test CI: -1.59, 22.71 *	T-test CI: -14.48, 12.62
Education level of father (schooling years)	Pearson’s r= -.003	Pearson’s r= .006
Income per capita (in MZM)	Pearson’s r= .071	Pearson’s r= -.245 (
Cooking method used [vapour/boiling]	T-test CI: -13.29, 11.22	T-test CI: -13.29, 11.22
Recipe chosen [5 recipes]	Spearman’s r= -.104	Spearman’s r= -.084
Number of ingredients used [2, 3 or 4]	Spearman’s r= -.054	Spearman’s r= -.167
Time interval (in days)	Pearson’s r= -.111	Pearson’s r= -.187

*p<0.1 **p<0.05 ***p<0.01 ****p<0.001

There was no association between recall scores and socio-demographic variables (table 5.22). There was also no association between recall scores and other exposure variables (those measured with the recipe task) (table 5.22).

5.5.2 Comparing recall scores between approach/intervention group

Table 5.23 Descriptive statistics for recall scores

	INTERVENTION GROUP	N	Mean	Std. Deviation
Recall scores	INTERVENTION I	40	72.84	19.97
	INTERVENTION II	40	80.04	19.50

Above table 5.23 and figure 5.10 compare recall mean scores between the two intervention groups. The mean unadjusted score was almost 10% higher for intervention II.

Figure 5.10 Comparing recall scores between intervention groups

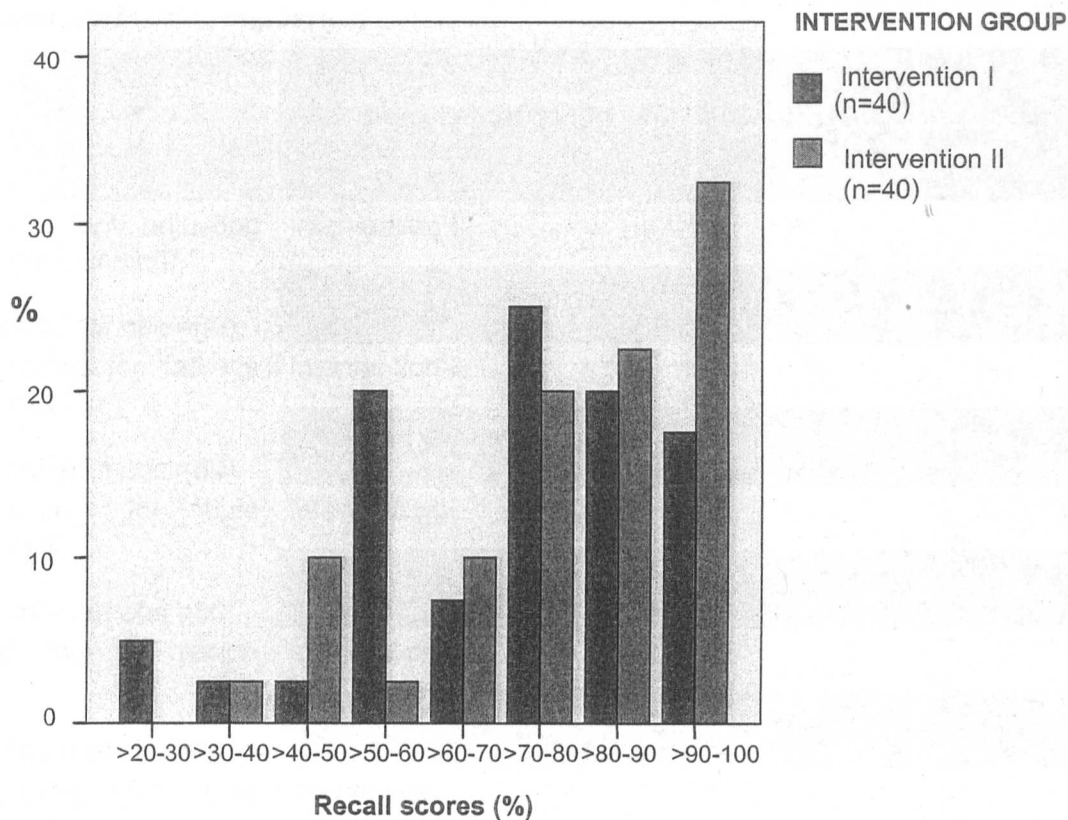


Table 5.24 explores the association between recall scores in intervention I with those in intervention II before and after adjusting for confounding.

After adjusting for father's education, the coefficient increased slightly, also increasing the statistical difference. After adjusting for father's education, scores for intervention II were marginally but non-significantly higher than intervention I. After adjusting for time interval, the coefficient decreased for intervention II, and the statistical difference remained non significant. Recipe type did not seem to be confounding the relationship since the coefficient did not change much. After adjusting for both father's education and time interval, the two potential confounders, the coefficient for intervention II decreased only slightly.

Table 5.24 Adjusting the association between intervention group and recall scores for confounders (using regression)

Model	Intervention group	Coefficient	P> z
Crude (not adjusting for confounding) (Number of obs =80)	Intervention I	1	"
	Intervention II	7.20	0.107
Adjusting for father's education (Number of obs =80)	Intervention I	1	
	Intervention II	8.24	0.086
Adjusting for time interval (Number of obs =80)	Intervention I	1	
	Intervention II	5.8	0.202
Adjusting for recipe type (Number of obs =80)	Intervention I	1	
	Intervention II	7.11	0.112
Adjusting for time interval and father's education (Number of obs =80)	Intervention I	1	
	Intervention II	6.64	0.175

5.6 Attitudes

In this section, attitudes results are compared between the intervention and control groups in two ways: by comparing each attitude item and by comparing the total score for each section. The control group consists of those women selected at random from the TSNI control group. It is included here as a proxy for baseline attitudes since baseline attitudes for the intervention groups were not collected.

P-P plots (not shown) for the cumulative proportion of intervention I and intervention II summed attitude scores against the cumulative proportion of a normal distribution have shown points clustering around the normal distribution straight line.

Table 5.25 compares mean values of several socio-demographic variables among the two intervention groups and the control group. Only education level of mother differed significantly between the two intervention groups. All other socio-demographic

variables did not differ among the 3 groups. Income per capita did not differ significantly between the two intervention groups (T-test, mean diff 134.15, CI: -114.34 – 95.98. $p=0.864$).

Table 5.25 Analysis of Variance for socio-demographic variables

SOCIO DEMOGRAPHIC VARIABLE	TREATMENT GROUP	Mean Difference (95% CI)	Sig.
AGE OF MOTHER (YEARS)	INTERVENTION I INTERVENTION II	-0.355 (-2.35, 1.64)	1.000
	INTERVENTION I CONTROL	-0.305 (-2.34, 1.73)	1.000
	INTERVENTION II CONTROL	-0.050 (-2.08, 1.98)	1.000
MOTHER - SCHOLASTIC LEVEL (YEARS OF SCHOOLING)	INTERVENTION I INTERVENTION II	-0.542 (-1.07, -0.02)	.040
	INTERVENTION I CONTROL	-0.263 (-0.80, 0.27)	.717
	INTERVENTION II CONTROL	0.279 (-0.25, 0.81)	.624
MAN - SCHOLASTIC LEVEL (YEARS OF SCHOOLING)	INTERVENTION I INTERVENTION II	-0.120 (-0.94, 0.70)	1.000
	INTERVENTION I CONTROL	-0.051 (-0.89, 0.79)	1.000
	INTERVENTION II CONTROL	0.069 (-0.77, 0.91)	1.000

Bonferroni correction included

Test-retest reliability

A retest reliability analysis on 10 subjects showed that scores were very stable over time for both section A (Pearson's correlation $r=0.862$, $p=0.001$) and section B (Pearson's correlation $r=0.771$, $p=0.009$). Cronbach's alpha which is designed to measure the internal consistency (i.e. "do all the items within each section measure the same thing?") was also good (Alpha, $r=0.702$) for section A when scores for item 5 and 7 were reversed, and was borderline for section B (Alpha, $r=0.698$).

5.6.1 Association between attitude sum scores and exposure factors

Socio-demographic variables were associated with attitudes scores in each intervention group (table 5.26).

Table 5.26 Association between attitude scores and socio demographic variables by intervention groups (using Pearson's r and Independent T-test (CI of the difference in means))

SOCIO DEMOGRAPHIC VARIABLES	BREASTFEEDING ATTITUDE SCORES		COMPLEMENTARY FEEDING ATTITUDE SCORES	
	INTERVENTION I	INTERVENTION II	INTERVENTION I	INTERVENTION II
Age of Mother (in years)	Pearson's r= .135 (n=139)	Pearson's r= .029 (n=142)	Pearson's r= .102 (n=139)	Pearson's r= .014 (n=142)
Education level of mother (schooling years)	Pearson's r= .095 (n=139)	Pearson's r= .048 (n=142)	Pearson's r= .028 (n=139)	Pearson's r= .347**** (n=142)
Mother has formal education? (Yes/No)	CI: -3.66, 1.56 (n=139)	CI: 3.95, 1.13 (n=142)	CI: 1.99, 2.85 (n=139)	CI: -5.62, -1.99**** (n=142)
Education level of father (schooling years)	Pearson's r= .166 * (n=118)	Pearson's r= .155* (n=117)	Pearson's r= .174 ** (n=118)	Pearson's r= .313*** (n=117)
Income per capita (in MZM)	Pearson's r= .116 (n=138)	Pearson's r= .080 (n=141)	Pearson's r= .017 (n=138)	Pearson's r= .115 (n=141)

*p<0.1 **p<0.05 ***p<0.01 ****p<0.001

Education level of the mothers in intervention II showed a positive medium association with their attitude scores for the complementary feeding section. This suggests that the higher the education level of the mother the more positive the attitudes. Education level of intervention II father's also showed a positive medium association with the mother's attitude scores for the complementary feeding section. This suggests that the higher the education level of the father the more positive the attitudes of the mother.

There were 4 health promoters working for the TSNI project which were allocated to different localities and it is possible that attitudes in one locality were higher than in another. Analysis of Variance was carried out in order to explore the associations between attitude scores for each section and health promoters.

Results from the ANOVA (not shown) suggest that there is an association between health promoters and attitude scores. Section A attitude scores' in Malei locality were significantly higher than those in Catale and Luala/Posto Campo (CI: 4.38 – 11.47, p<0.001; and CI: 0.41 – 6.75, p=0.018, respectively). Attitude scores for section A in Mexixine locality were significantly higher than those in Catale (CI: 1.85 – 8.73,

$p < 0.001$) but not different from those in Malei and Lua/Posto Campo. Attitude scores for section B in Malei, Lua/Posto Campo and Mexixine localities were significantly higher than those in Catale (CI: 4.17 – 10.35, $p < 0.001$; and CI: 2.90 – 8.38, $p < 0.001$; and CI: 6.10 – 11.58, $p < 0.001$, respectively). Attitude scores for section B in Mexixine locality were significantly higher than those in Lua/Posto Campo (CI: 0.77 – 5.63, $p = 0.003$) but not different from those in Malei. There was also no difference between Malei and Lua/Posto Campo.

Even though the characteristics of each health promoter was not measured *per se* it was noted that the health promoter in Malei was the most motivated and with more initiative. Her spirit and her energy in the work probably led her to be more persuasive than the others.

5.6.2 Attitudes towards breastfeeding: comparison between the two intervention groups and with the control group

Table 5.27 shows the proportions of answers to each attitude item in the breastfeeding section of the attitude tool by intervention group.

When the proportion of positive attitudes for each item was pooled together and averaged for both interventions, it was possible to observe that five items showed quite high positive attitudes, ranging from 74% to 90%, but three items: 2 (55.5%), 9 (48%) and especially item 6 (33.4%) showed lower proportions of positive attitudes. Item 7 was used as a bogus to catch inattentive mothers or those mothers that tended to agree or disagree with all the items. The expected response to item 7, as tested, would be for mothers to agree with it. The observed response, as shown in the table, was that 98.6% in intervention I and 95% in intervention II agreed with it. So did, 97% in the control area.

Appendix 4.7 describes the reasons given by mothers that agreed with the wrong breast feeding practice.

Table 5.27 also presents the results of the comparisons of each item's response among intervention and control groups. As can be seen from the chi-square tests comparing the two intervention groups, at the end of the TSNI study, attitudes towards breastfeeding practices were very similar for both intervention groups. When using the control group as a proxy for baseline attitudes, it can be seen that both intervention

groups had significantly higher positive attitudes towards breastfeeding practices than the control group.

5.6.3 Attitudes towards complementary feeding: comparison between the two intervention groups and with the control group

Table 5.28 shows the proportions of answers to each attitude item in the complementary feeding section of the attitude tool by intervention group.

When the proportions of positive attitudes for each item was pooled together and averaged for both interventions, it was possible to observe that five items (i.e. 11, 12, 14, 16, 18) showed quite high positive attitudes, ranging from 68% to 99%, while two items: 10 (51%) and especially 15 (16.7%) showed lower proportions of positive attitudes.

Table 5.28 also presents the results for the comparisons of each item's response between intervention groups and between each of the intervention groups and control group.

Whilst scores for most items in section A were significantly higher for both intervention groups when compared to the control group, the same did not happen with scores for section B. Intervention II seemed to score better than intervention I. Significantly higher positive attitudes for intervention II mothers were only seen for 4 (i.e. 10, 11, 13, 17) items. For all the other items, there was no significance difference, except with item 15, where scores were significantly higher for the control group.

Table 5.27 Comparing proportion of answers to attitude questions regarding breast feeding practices between the two intervention groups (intervention I n=139, intervention II n=142 and control n=130)

Likert Scale ¹	1 "Remove the first milk and dispose of it"		2 "Give herbal tea to the new born baby"		3 "Decrease breast feeding and it will decrease diarrhoea"		4 "Stop breastfeeding at 6 months to be able to go to the field"		5 "Continue breastfeeding even if the mother is a month pregnant"		6 "Stop breastfeeding if the mother is 4 months pregnant"		7 "Continue breastfeeding a 3 month infant even if the mother is ill with malaria"		8 "Stop breastfeeding a 9 month baby if the mother is ill with malaria"		9 "Stop breastfeeding if mother has not breastfed for 3 days"	
	Interv I	Interv II	Interv I	Interv II	Interv I	Interv II	Interv I	Interv II	Interv I	Interv II	Interv I	Interv II	Int I	Int II	Int I	Int II	Int I	Int II
1	61.9%	61.3%	44.6%	41.5%	57.6%	52.8%	63.3%	57.7%	13.7%	13.4%	26.6%	24.6%	1.4%	3.5%	74.1%	69.0%	43.2%	27.5%
2	8.6%	14.8%	10.1%	14.8%	16.5%	23.9%	9.4%	18.3%	2.9%	9.2%	7.9%	7.7%	.0%	.7%	15.8%	20.4%	10.8%	16.2%
3	0	0	0	0	0	0	0	0	0	0	0	0	.0%	.7%	.0%	.7%	.7%	.7%
4	2.9%	.7%	7.2%	14.8%	10.8%	7.0%	7.2%	4.2%	4.3%	2.1%	5.8%	5.6%	6.5%	4.2%	2.2%	1.4%	2.2%	7.0%
5	26.6%	23.2%	38.1%	28.9%	15.1%	16.2%	20.1%	19.7%	79.1%	75.4%	59.7%	62.0%	92.1%	90.8%	7.9%	8.5%	43.2%	48.6%
Chi-x ² (interv)	p=0.293		p=0.293		p=0.605		p=0.515		p=0.206		p=0.704		p=0.157		p=0.968		p=0.083	
Odds ratio	1.329		1.329		1.154		1.195		1.467		0.908		3.044		1.016		0.659	
Chi-x ² (control)	p<0.001 ³		p<0.001		p<0.001		p<0.001		p<0.001		p<0.001		p=0.358		p<0.001		p<0.001	
Strongly disagree = 1; Disagree = 2; Indecisive = 3; Agree = 4; Strongly Disagree = 5; Chi- square test – intervention group against positive attitudes (1 + 2) and negative attitudes (4 + 5)																		
Chi-x ² (interv) shows results for the comparison between intervention groups; Odds ratio for the comparison between intervention groups; Chi-x ² (control) shows results for the comparison between intervention groups and the control group; Answers=indecisive (3) have been omitted; ³ same p value for the comparison with both intervention groups																		

Table 5.28 Comparing proportion of answers to attitude questions regarding complementary feeding practices between the two intervention groups (intervention I n=139, intervention II n=142 and control n=130)

Likert Scale ¹	10 "Enough for a crawling baby to eat twice a day"		11 "6 month infant should not eat egg"		12 "Should not insist the child to eat if he/she refuses to"		13 "It is good for the baby to eat porridges made of cassava flour and sugar"		14 "It is okay for a 2 yr old child to skip breakfast"		15 "It is very important for the child to eat fat-rich foods and 3 times a week is enough"		16 "A 2-yr old child is old enough to eat by herself without the mother's supervision"		17 "It is okay for a 2-yr old child to share the food with the parents"		18 "There is not enough liver so it should be given to the father"	
	Interv I	Interv II	Interv I	Interv II	Interv I	Interv II	Interv I	Interv II	Interv I	Interv II	Interv I	Interv II	Int I	Int II	Int I	Int II	Int I	Int II
1	34.5%	41.5%	78.4%	70.4%	73.4%	63.4%	56.1%	55.6%	68.3%	66.9%	12.2%	9.2%	51.1%	57.0%	43.2%	61.3%	80.6%	80.3%
2	10.8%	15.5%	12.9%	18.3%	24.5%	32.4%	11.5%	13.4%	18.0%	29.6%	5.0%	7.0%	24.5%	25.4%	15.1%	19.7%	10.1%	15.5%
3	.0%	.7%	0	0	0	0	1.4%	.7%	0	0	0	0	.7%	.0%	0	0	0	0
4	8.6%	14.1%	1.4%	4.2%	.0%	1.4%	4.3%	4.9%	.7%	1.4%	4.3%	4.9%	7.9%	4.2%	5.0%	3.5%	0	0
5	46.0%	28.2%	7.2%	7.0%	2.2%	2.8%	26.6%	25.4%	12.9%	2.1%	78.4%	78.9%	15.8%	13.4%	36.7%	15.5%	9.4%	4.2%
Chi-x ²	p=0.042		p=0.461		p=0.325		p=0.872		p=0.002		p=0.810		p=0.193		p<0.001		p=0.087	
Odds ratio	1.629		1.344		0.500		1.043		4.338		0.926		1.471		3.050		2.339	
Chi-x ² (control)	p>0.05/ p<0.05 ³		p<0.001		p=0.162		p=0.086		p=0.114		p=0.014		p=0.893		p=0.677		p=0.328	
¹ Strongly disagree = 1; Disagree = 2; Indecisive = 3; Agree = 4; Strongly Disagree = 5 ³ Chi-square p value for item10 control/interv1 = 0.618, and for control/interv2 = 0.028																		

5.6.4 Comparing summed scores for each section between the intervention groups and control group

Figure 5.11 depicts the distribution of the sums of scores for the breastfeeding section for each group.

The following table 5.29 depicts the descriptive statistics for breastfeeding section scores by group. It should be noted that because of reverse scoring, the higher the mean, the higher the positive attitudes. Taken this into consideration it can be seen that while the mean between the intervention groups was very similar, the mean for the control group was lower.

Figure 5.11 Comparing distribution of sum of breastfeeding- attitude scores among groups

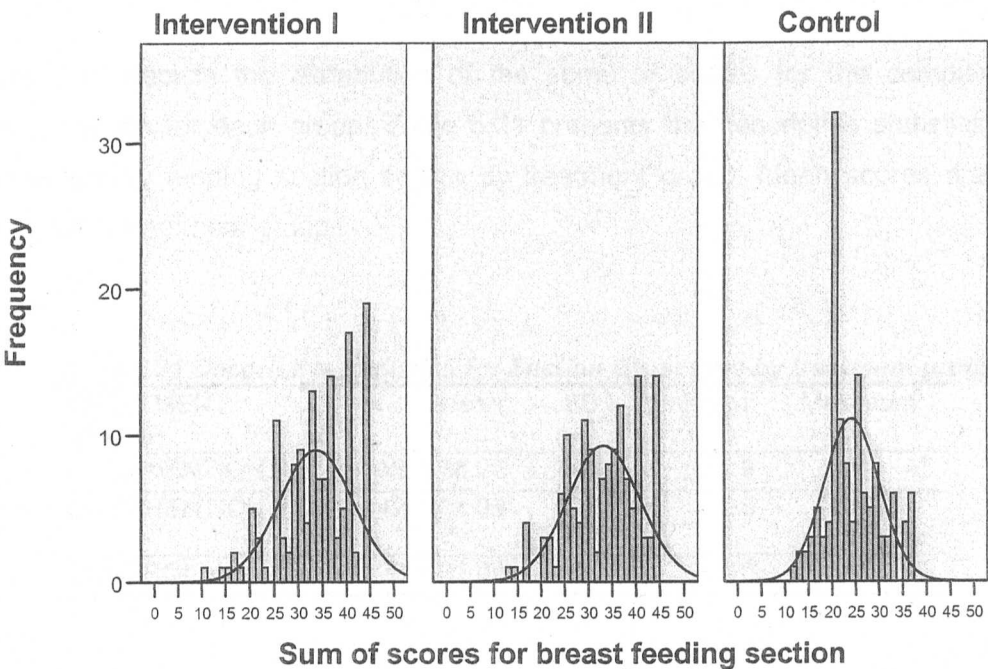


Table 5.29 Descriptive statistics for Section A's scores for each treatment group

TREATMENT GROUP	N	Mean	SD	Minimum	Maximum
INTERVENTION I	139	33.82	7.74	11	45
INTERVENTION II	142	33.25	7.64	13	45
CONTROL	130	24.19	5.86	12	37

Table 5.30 gives a detailed description of the differences between each of the groups using an ANOVA test. There was no significant difference between the two intervention groups. However, both intervention I and II mean scores were significantly higher than the control group.

Table 5.30 Analysis of Variance for the sum of scores - breastfeeding attitudes (Bonferroni correction included)

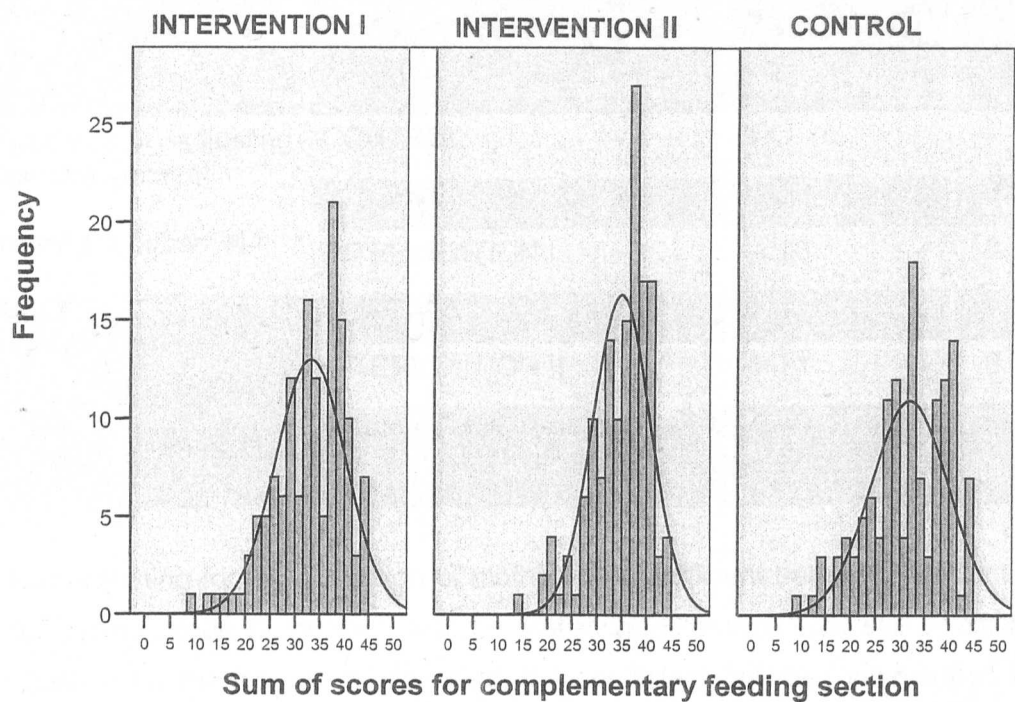
Treatment group	Mean Difference (95% CI)	Sig.
Intervention I	0.567 (-1.49, 2.62)	1.000
Intervention II		
Intervention I	9.628 (7.53, 11.73)	<0.001
Control		
Intervention II	9.061 (6.97, 11.15)	1.000
Control		

Figure 5.12 depicts the distribution of the sums of scores for the complementary feeding section for each group. Table 5.31 presents the descriptive statistics for the complementary feeding section scores by treatment group. Mean scores are similar among the 3 treatment groups.

Table 5.31 Descriptive statistics for Section B's scores by treatment group

TREATMENT GROUP	N	Mean	SD	Minimum	Maximum
INTERVENTION I	139	33.25	7.13	9	45
INTERVENTION II	142	34.09	5.79	15	45
CONTROL	130	32.02	7.90	10	45

Figure 5.12 Comparing distribution of sum of complementary feeding-attitude scores among treatment groups



As seen previously, there is an association between intervention group and education level of mothers and an association between intervention II scores for the complementary feeding section and education level of mother. Education level could thus be a confounder for the association between intervention group and attitude scores for the complementary feeding section. Table 5.32 shows the association between treatment group (i.e. intervention and control groups) and attitude scores, before and after adjusting for mother's education level.

Table 5.32 Adjusting the association between attitude scores and treatment group (multiple regression)

Model	TREATMENT GROUP	Coefficient	P> z
Crude (not adjusting for confounding) Number of obs= 411	CONTROL	1	
	INTERVENTION II	-2.93	0.001
	INTERVENTION I	-1.23	0.147
Adjusting mother's education level Number of obs= 411	CONTROL	1	
	INTERVENTION II	-2.71	0.001
	INTERVENTION I	-1.43	0.087

After adjusting for the education of mother, the coefficient between intervention II and the control group decreased, but the statistical difference was still significant. After adjusting for the education of mother, the coefficient between intervention I and the control group increased, and the statistical difference increased accordingly.

Table 5.33 shows the association between the two intervention groups and attitude scores, before and after adjusting for mother's education level.

Table 5.33 Adjusting the association between intervention groups and attitude scores (multiple regression)

Model	TREATMENT GROUP	Coefficient	P> z
Crude (not adjusting for confounding) Number of obs= 411	INTERVENTION I	1	
	INTERVENTION II	1.69	0.030
Adjusting mother's education level Number of obs= 411	INTERVENTION I	1	
	INTERVENTION II	-1.31	0.089

After adjusting for the education of mother, the coefficient between intervention I and intervention II became non-significant.

5.6.5 Comparing section A to section B for each group

When sum of scores was compared between the sections, it was observed that there was no significant difference in the summed scores between section A and B for intervention I, there was a significant lower score (paired t-test, CI: -2.85, -0.53; $p=0.005$) in section A (i.e. more negative attitudes for section A) for intervention II and also a significantly lower score (paired t-test, CI: -9.02, -6.62; $p<0.001$) in section A for the control group.

5.7 Associations between the output steps

Because McGuire himself has discussed the relationship between the output steps and another study[78] has also looked at it, this section explores the association between the output steps.

The following table shows a matrix depicting the various association among all the output steps. Associations were carried out using Pearson's r unless stated otherwise.

Table 5.34 Associations between output steps for each approach-intervention group

GC Intervention I	Attendance level	Attention (obs)	Attention (recall-questions)	Comprehension Lexical-meaning	Comprehension Psychological-meaning	Skill acquisition	Recall	Attitude Section A
Attendance levels					r= -0.045 (n= 40)	r= -0.023 (n= 40)	r= 0.012 (n= 40)	r= 0.026 (n=139)
Attention (observation)			T-test CI:-1.91--1.10 (n=41)	T-test CI:-2.01 - .80 (n=41)				
Attention (recall-questions)		T-test CI:-1.91- 1.10 (n=41)		Chi-x ² , p=0.509 (n=41)				
Comprehension Lexical-meaning		T-test CI:-2.01 - .80 (n=41)	Chi-x ² , p=0.509 (n=41)					
Comprehension Psychological-meaning	r= -0.045 (n= 40)					r= 0.146 (n=40)	r=0.502**** (n=40)	r= 0.155 (n=28)
Skill acquisition	r= -0.023 (n= 40)				r= 0.146 (n=40)		r=0.335 ** (n=40)	r= 0.404 ** (n=28)
Recall	r= 0.012 (n= 40)				r=0.502 **** (n=40)	r=0.335** (n=40)		r= 0.163 (n=28)
Attitude Section B	r=0.088 (n=139)				r= 0.239 (n=28)	r= 0.395 ** (n=28)	r= 0.035 (n=28)	

- ID of mothers not known therefore not possible to associate them with other output steps

- association r = 1

Block 3 mothers in GC intervention II are the same as IC intervention II explaining why results for associations among block 3 output steps give the same results for both approach-intervention groups. *p<0.10 ** p<0.05 *** p<0.01 ****p<0.001

CI = 95 % confidence interval of the difference in the mean

GC Intervention II	Attendance levels	Attention (obs)	Attention (recall questions)	Comprehension Lexical-meaning	Comprehension Psychological-meaning	Skill acquisition	Recall	Attitude Section A
Attendance levels					r=0.126 (n=38)	r=-0.201 (n=38)	r=0.089 (n=38)	r=.231 *** (n=142)
Attention (observation)			T-test CI:-1.718 -- .547 (n=57)	T-test *** CI:-2.265 -- -0.219 (n=57)				
Attention (recall- questions)		T-test CI:-1.718-- 0.547 (n=57)		Chi-X ² ** (n=57)				
Comprehension Lexical-meaning		T-test ** CI:-2.265 -- 0.219 (n=57)	Chi-X ² ** (n=57)					
Comprehension Psychological-meaning	r=0.126 (n=38)					r=0.332 ** (n=40)	r=0.345 ** (n=40)	r=0.472 ** (n=19)
Skill acquisition	r=-0.201 (n=38)				r=0.332 ** (n=40)		r=0.365 ** (n=40)	r=0.353 (n=19)
Recall	r=0.089 (n=38)				r=0.345 ** (n=40)	r=0.365 ** (n=40)		r=-0.252 (n=19)
Attitude Section B	r=0.105 (n=142)				r=0.241 (n=19)	r=0.086 (n=19)	r=0.0971 (n=19)	

- ID of mothers not known therefore not possible to associate them with other output steps
 Block 3 mothers in GC intervention II are the same as IC intervention II explaining why results for associations among block 3 output steps give the same results for both approach-
 intervention groups. *p<0.10 ** p<0.05 *** p<0.01 **** p<0.001 CI = 95 % confidence interval of the difference in the mean

IC Intervention II	Attendance levels	Attention (obs)	Attention (recall questions)	Comprehension Lexical-meaning	Comprehension Psychological-meaning	Skill acquisition	Recall	Attitude Section A
Attendance levels		r= -0.005 (n=74)	T-test Cl: -.718-0.289 (n=74)	T-test Cl: -0.473 - 0.194 (n=73)	NM	NM	NM	r= 0.161 ** (n=141)
Attention (observation)	r= -0.005 (n=74)		T-test Cl: -0.339 - 1.906 (n=91)	T-test Cl: -0.449 - 1.372 (n=91)	r= -0.375 (n=20)	r= -0.191 (n=20)	r= -0.160 (n=20)	r=0.164 (n=43)
Attention (recall- questions)	T-test Cl: -.718 - .289 (n=74)	T-test Cl: -.339 - 1.906 (n=91)		Chi-x ² , p=0.610 (n=91)	T-test Cl: -29.5 - 29.5 (n=20)	T-test * Cl: -.15 - 1.88 (n=20)	T-test Cl: -17.9 - 14.1 (n=20)	T-test Cl: -6.6 - 4.8 (n=43)
Comprehension Lexical-meaning	T-test Cl: 0.473 - 0.194 (n=73)	T-test Cl: -0.449 - 1.372 (n=91)	Chi-x ² , p=0.610 (n=91)		T-test Cl: -48.80 - 17.55 (n=20)	T-test Cl: -2.28 - 1.40 (n=20)	T-test Cl: -12.94 - 39.73 (n=20)	T-test Cl: -3.3 - 10.5 (n=43)
Comprehension Psychological-meaning	NM	r= -0.375 (n=20)	T-test Cl: -29.5 - 29.5 (n=20)			r=0.332 ** (n=40)	r=0.345 ** (n=40)	r= 0.472 ** (n=19)
Skill acquisition	NM	r= -0.191 (n=20)	T-test * Cl: -.15 - 1.88 (n=20)	T-test Cl: -2.28 - 1.40 (n=20)	r=0.332 * (n=40)		r=0.365 ** (n=40)	r= 0.353 (n=19)
Recall	NM	r= -0.160 (n=20)	T-test Cl: -17.9 - 14.1 (n=20)	T-test Cl: -12.94 - 39.73 (n=20)	r=0.345 ** (n=40)	r=0.365 ** (n=40)		r= 0.252 (n=19)
Attitude Section B	r= 0.072 (n=141)	r= 0.058 (n=43)	T-test * Cl: -5.3 - 0.5 (n=43)	T-test Cl: -2.5 - 5.6 (n=43)	r= 0.241 (n=19)	r= 0.086 (n=19)	r= 0.0971 (n=19)	

- ID of mothers not known therefore not possible to associate them with other output steps - association $r = 1$

Block 3 mothers in GC intervention II are the same as IC intervention II explaining why results for associations among block 3 output steps give the same results for both approach-intervention groups. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$ **** $p < 0.001$ CI = 95 % confidence interval of the difference in the mean

Exposure/attendance levels and other output steps:

Exposure and block 2 output steps: there is no association between attendance levels to IC with attention scores (measured with either method) and comprehension lexical meaning scores measured in IC. However, this lack of association could be due to the fact that the group of mothers that did not attend all 6 IC sessions consisted only of 2 women, leaving not much room for variation.

Exposure and block 3 output steps: There is no association between Intervention I attendance levels in GC and any of the block 3 output steps either presented as continuous variables or when they were categorised into two categories (i.e. 0-50% and above 50%).

There is also no association between Intervention II attendance levels in GC and any of the block 3 output steps either presented as continuous variables or when they were categorised into two categories. It was not possible to associate attendance to IC and block 3 output steps as all mothers that fell in both groups had perfect attendance.

Exposure and attitudes for section A and B: As can be seen from the matrix, there was no association between attitudes in either section and Intervention I mother's attendance levels to GC.

Intervention II mothers attendance levels to both GC and IC had a negative association with attitudes for section A (hence, the more classes they attended the more positive their attitudes). There was no association with section B.

Block 2 output steps (i.e. attention observation, attention recall and understanding-lexical meaning) and other output steps:

There was no association between the two attention methods (i.e. observation and recall questions) when it was looked separately by approach-intervention group.

There was no association between attention scores, measured with either method, and comprehension lexical meaning for both GC intervention I and IC. However, for GC intervention II there was a positive association between attention observation and comprehension lexical meaning with a higher attention observation mean for the group that answered the comprehension question correctly. For GC intervention II there was also a positive association between attention recall and comprehension lexical

meaning with a higher attention recall for the group that answered the comprehension question correctly.

Block 2 output steps and block 3 output steps: there is no association among the two blocks of output steps.

Block 3 output steps (i.e. comprehension-psychological meaning, skills acquisition, and recall):

As can be seen from the matrix, apart from one non-significant association between skills acquisition and comprehension for intervention I mothers, all other associations were significant.

Attitude:

Attitudes and block 2 outcomes: there was no association between attitudes in either section and Intervention II mother's attention levels (measured with both methods) and comprehension-lexical meaning level in IC.

Attitudes and block 3 outcomes: There was no association between attitudes from both sections and comprehension-psychological meaning and recall levels measured in both intervention I and II mothers.

There was also no association between attitudes from both sections and skills acquisition levels measured in intervention II mothers. But there was, however, a negative association between attitudes from both sections and skills acquisition levels measured in intervention I mothers.

5.8 Comparing persuasiveness between the two Intervention groups

Finally, to determine the overall impact of the delivery method (i.e. intervention I and II), a measure of persuasiveness for each intervention group was calculated. This consisted of weighting each output step (table 5.35) and then creating an overall score of persuasiveness using these weights (see section 4.8.5). The mean of this score was compared between the intervention groups.

For attendance levels, where measurements were carried out in the same subjects for both GC and IC, an average of the two was first calculated before adding the weights.

Table 5.35 *Weights allocated to each output step*

Output steps	Attendance levels	Attention (obs)	Attention (recall)	Comprehension Lexical-meaning	Comprehension Psychological-meaning	Attitude section A	Attitude section B	Recall	Skill acquisition
Weights	0.25	0.35	0.35	0.35	0.40	0.50	0.50	0.55	0.60

The following table 5.36 presents the persuasiveness characteristics broken down by output step for each intervention group. Mean weighted scores in each output step were always higher for intervention II except with attitudes for section A where it was higher for intervention I.

When an overall score was created (i.e. when the weighted scores of all outcomes were pooled for each intervention group) and then averaged to reflect a factor of persuasiveness, the persuasiveness degree in intervention II was significantly higher than that for intervention I ($p=0.006$).

Table 5.36 Descriptive statistics of the weighted scores for each output step by intervention group and comparison of weighted scores between the intervention groups for each output step (using T-test)

Output step	Intervention group	N	Minimum ¹	Maximum ¹	Mean (SD)	95% CI	p-value
Attendance levels	I	274	.00	25.00	18.02 (6.73)	-2.43, -.28	0.013
	II	286	.00	25.00	19.38 (6.18)		
	Total	560	.00	25.00	18.71 (6.48)		
Attention obs	1	42	7.00	35.00	26.33 (7.62)	-4.35, 0.92	0.199
	2	148	.00	35.00	28.05 (7.27)		
	Total	190	.00	35.00	27.67 (7.36)		
Attention recall	I	42	.00	35.00	27.50 (9.58)	-6.56, -0.15	0.040
	II	148	17.50	35.00	30.86 (7.46)		
	Total	190	.00	35.00	30.12 (8.07)		
Comprehe nsion ²	I	42	.00	35.00	19.16 (17.63)	-11.0, 1.10	0.107
	II	148	.00	35.00	24.12 (16.25)		
	Total	190	.00	35.00	23.02 (16.65)		
Comprehe nsion ³	I	40	.00	40.00	21.99 (11.76)	-7.05, 2.70	0.377
	II	40	5.00	40.00	24.17 (10.08)		
	Total	80	.00	40.00	23.08 (10.94)		
Skills	I	40	20.00	60.00	41.12 (12.17)	-15.27, -5.22	<0.001
	II	40	20.00	60.00	51.37 (10.31)		
	Total	80	20.00	60.00	46.25 (12.34)		
Recall	I	40	13.75	55.00	40.06 (10.98)	-8.79, 0.87	0.107
	II	40	18.33	55.00	44.02 (10.72)		
	Total	80	13.75	55.00	42.04 (10.97)		
Attitude A	I	139	2.78	50.00	34.47 (10.74)	-1.72, 3.24	0.537
	II	142	5.56	50.00	33.68 (10.61)		
	Total	281	2.78	50.00	34.07 (10.66)		
Attitude B	I	139	.00	50.00	33.68 (9.91)	-4.4, -0.23	0.030
	II	142	8.33	50.00	36.03 (8.04)		
	Total	281	.00	50.00	34.87 (9.07)		
Total	I	798	.00	60.00	27.07 (12.82)	-2.74, -0.46	<u>0.006</u>
	II	1134	.00	60.00	28.67 (12.39)		
	Total	1932	.00	60.00	28.01 (12.59)		

¹ Values for all output steps are proportions ² Comprehension-lexical meaning ³ Comprehension psychological meaning

CHAPTER 6

DISCUSSION AND CONCLUSIONS

This concluding chapter discusses the results from the previous results chapter whilst also addressing the research questions:

- Can McGuire's Communication/Persuasion model detect differences in the communication process between the two nutrition education interventions?
- Were the tools developed able to detect variability within groups for each of the output steps?
- Do exposure, attention, understanding, skills acquisition, recall, and attitude scores differ significantly between the intervention groups?
- Does the degree of persuasiveness differ significantly between the intervention groups?

This chapter is divided into 3 main sections: Section 1 discusses the use of McGuire's model as a conceptual framework for the process evaluation of the TSNi nutrition education program. Section 2 consists of a critique of the tools developed by the researcher for this study and which were used for measuring each of the output steps. Section 3 discusses the results for the comparison of each output step between the two intervention groups. For each section, limitations are addressed and improvements/recommendations for future research are suggested. Main study design limitations are then addressed. The chapter ends by summarising the main findings and their possible implications.

Health behaviour involves the subtle interplay of human beings interacting in a rich, complex environment of which the persuasive message is only a small part. While this research study only examines the impact of the delivery method on the success of the TSNi nutrition education program through the evaluation of McGuire's output steps, it does not exclude the possible impacts that external factors (e.g. traditional beliefs and social pressure) might have on the persuasiveness of the communication. Any conclusions from this study would therefore be strengthened by an in-depth qualitative study of the impact of cultural and socio factors on the success of the TSNi nutrition education program. Moreover, the communication process itself could have been evaluated using qualitative methodologies, for example, by conducting interviews such as in-depth interviews, or semi structured interviews, as well as focus group

discussions with the women, and direct observations. Using these methods, one could explore the following: what drove them to attend the classes; if they missed a lot of classes, then what were the reasons; if they enjoyed the classes, what classes did they enjoy the most, why did they enjoy these the most; if they relate with the topics given; if they found the topics useful; if they found it difficult to concentrate in the classes or maybe some topics; if they found the material difficult to understand and what made it difficult to understand and what could be done to make it easier to understand; if they used the information learned in any occasion, like for example to tell it to someone or maybe to change a dietary practice and were able to retrieve the information learned whenever they needed; do they feel that they are capable to apply the information they learned and if not then why. Using these methods one could explore attitudes and could also confirm if behaviours have changed. Observations would probably be the best method to measure behaviour change. The TSNI study carried out direct observations during lunch time in intervention II households where the health promoters observed dietary practices as well as hygiene practices.

6.1 McGuire's model and persuasive communication

A critical appraisal of McGuire's model is the focus of this section, this includes a critical appraisal of the sequence of output steps and how useful they are as tools for process evaluation of nutrition education programs.

This research study included a systematic review of those studies that also adopted McGuire's output steps in order to evaluate a communication. This section will discuss this research study's results while referring and comparing to the results of some of these others studies.

6.1.1 Understanding the communication process

As was discussed in the beginning of this thesis, it is imperative that we understand the communication process which explains why a health message leads (or does not lead) to a person changing her/his behaviour. McGuire's communication/persuasion model was selected in this research study to allow for the understanding of this communication process.

In this research study, McGuire's model allowed for the breakdown of the communication persuasion process into steps which gave an idea of the effect of the

message in several steps along the communication process. This in turn shows the researcher in which steps the effect of the message is weaker or stronger.

6.1.2 User-friendliness of McGuire's model

The use of McGuire's model in other studies (used in systematic review carried out for this thesis) and the experience from this research study indicates that the model is a very flexible one and hence, user-friendly. This comes about due to its easiness to adjust and adapt to the study's objectives and setting; due to the fact that McGuire himself left the measurement of the output steps open to the researcher and thus open to the inclusion of innovative tools to measure the output steps. Moreover, it does not limit the way and number of output steps measured. It has been discussed that it is sometimes difficult to distinguish between the output steps in terms of how they are to be operationally defined, and McGuire himself has suggested that the division points between steps are arbitrary and some "simplifications" of the model can be made for empirical studies.[87] For these reasons, as was shown in chapter 2, studies differ in the number of output steps measured, and have also adopted different methods to measure the output steps. Some have only measured 3 output steps[84] and others have grouped together some output steps.[85]

However, other researchers might find that the flexibility of the model is not an acceptable quality. They may prefer a model that comes with the guidance about the measurement of the output steps, and where the constructs have been interpreted and measured similarly in other studies.

6.1.3 Delivery methods as input steps in the model

As discussed by McGuire, the aim of the input/output matrix should be to explore possible ways (i.e. manipulating input step characteristics) of maximising effectiveness at each output step.

Most of the studies referred to in the systematic review have also looked at the effects of a set of input step characteristics on several output steps. Bull *et al* (2001)[87] study for example looked at the effects of a range of input step characteristics (*receiver* factors – 4 psychosocial characteristics [eg.: readiness to lose weight] + 4 demographic and physical characteristics [eg.: age, gender, level of education, and BMI]; *channel* factors – individual preparedness to read printed materials; *source* factors – perceived attractiveness of the material and the degree they found it trustworthy; and *message*

factors – if participants found the material informative, encouraging, useful, applicable to their lives and tailored versus untailored messages) on 10 output steps: exposure, attention, liking, understanding, skills acquisition, yielding, memory storage/agreement, information search and retrieval, deciding and behaving.

Likewise, this research study has explored how to increase persuasiveness by manipulating one characteristic of the channel (i.e. delivery methods: intervention I and intervention II). Because this research study is only studying 1 characteristic of 1 input step (i.e. channel), one could argue that the observed results might be a result of other input step characteristics which were not measured in this research study. However, looking at the effects of a range of input step characteristics on output steps was not the scope of this research study, and indeed only Bull *et al*, and Russon and Koehly's (1995) [85] studies looked at a matrix of a whole range of input characteristics versus a range of output steps. All the other studies described in the systematic review, have only looked at 1 or 2 input characteristics.

Results varied a lot for all the studies. Zhao and Ramaprasad's (2005) study[89] for example, which looked at the effectiveness of different levels of fear appeals of advertisements on output steps, showed that higher fear appeals in HIV advertisements increase attention to it but do not increase interest, liking and intention to behave. Likewise, this research study has shown that there were no differences in recall and comprehension-psychological meaning between the two intervention groups but there were for all the other output steps. This is expected, as McGuire himself discusses that progression through the steps can stop anywhere in the sequence. Moreover, results for the various studies could not have been the same since they explored effects of different input steps. This research study is the only one that explores the effect of different channels of communication (i.e. the delivery methods). Results for the different effects that the various input steps have on the communication process are useful as they can help inform the health planner on what aspects of the health education/program might be improved. In the case of this research study, contributing with suggestions for future research on how the delivery methods can contribute to developing more effective health education/communication. McGuire's list of output steps points to additional ways in which the delivery methods could affect the communication/persuasion process of a health education program in a developing country setting, which has not yet been studied by health education researchers.

6.1.4 Exploring the model

The model also calls attention to possible interactions among the output steps. For example, the added benefit of receiving IC might mean that intervention II mothers perform better in most of the output steps, however, it also means that they attend less group classes which could have an impact on their acquired knowledge and even change in attitudes.

Still, this research study was not designed to try and explain interactions between the output steps, or test the model *per se*, exploring the linearity of the model or the progression through the output step sequence. This weakness in the study design derives from the fact that to be able to explore associations between the output steps and look at subjects' progression through the steps, measurements for all output steps would have to be taken for the same subjects. Moreover, because of this weakness in the study, behaviour change was also not measured and hence, it is not possible to look at how each output step could contribute or explain the variance in behaviour change.

Nonetheless, results for those associations carried out among the output steps showed that there was no association between attendance levels and all the other output steps. Hence, attending more classes had no effect on the scores of consecutive steps as would have been expected. However, assessment of this association was restricted by the sample size of the consecutive output steps, and lack of variability in IC attendance levels.

There was also no association between the two attention methods' results. Hence, higher scores for the direct observation method did not necessarily mean better recall of attention questions as would have been expected. Several justifications for this lack of association are given in the previous section. Attention observations might reflect visual attention while attention recall might reflect auditory attention. Or as suggested by Dr Casimir Ludwig from the University of Bristol, attention recall might involve a strong memory component which might mean that even though the mother is unable to remember something (i.e. the attention recall questions) it may be that she encoded it properly (i.e. sustained attention measured with the direct observation method), but is not able to retrieve it (personal communication).

There was also no association between either of the two attention methods' results and comprehension-lexical meaning for IC and GC intervention I, but there was a positive association in GC-intervention II, with higher attention observation scores and higher

recall for those mothers that correctly answered the understanding question. It is possible that the lack of association in GC intervention I was due to a small sample size restricting variability.

There was no association between block 2 output steps (i.e. attention observation, attention recall and comprehension-lexical meaning) and block 3 output steps (i.e. comprehension-psychological meaning, skills acquisition, and recall). However, this was expected, as first of all, measurement of this association was restricted to 20 IC mothers, a very small sample size to see any variation. But even if the sample size was much larger, there is the fact that attention was only measured once for each mother, reflecting their attention levels of only that day, while block 3 output steps reflect an acquisition of knowledge. Perhaps, to more accurately compare these blocks, future studies should consider measuring attention-recall and comprehension lexical-meaning together with block 3 output steps, hence, as part of the recipe task. It would, however, be very difficult to carry out direct attention observations in these recipe demonstrations as mothers are constantly moving around and in many occasion the researcher would not have an easy view of the subject's face.

There were strong positive associations among the block 3 output steps, hence for example, those mothers that scored highly in the skills acquisition task also scored highly for recall. What is not possible to say is which of the output steps accounted for the variability in the other(s). While McGuire believed that there was a linear relationship between the output steps, hence, comprehension was followed by skills acquisition and this in turn by recall, however, other studies have disagreed with this, and indeed Rimer and Glassman (1984)[78] discuss that "the path is clearly more complicated than a straight line connecting all the variables. While this study found clear connections among the output steps it also showed that they seem to have worked through and around one another." [78]

McGuire has positioned attitudes just after skills acquisition in his linear process, and in this study attitudes for both sections were positively associated with the skills acquisition of intervention I mothers, However, this could have been a coincidence since the same was not seen for intervention II mothers. Moreover, attitudes for section A were positively correlated with comprehension psychological meaning of intervention II mothers but not with that of intervention I mothers, suggesting that results were inconsistent. This might derive from the fact that associations were only carried out on 28 subjects which is probably a small sample size to detect any variability.

6.1.5 Future use of McGuire's model for evaluation

While output steps measured this way cannot say much about the linearity of the model and association between them cannot be assessed, they can still be used to carry out cross sectional measures of the performance of a nutrition education program. For example, they can be developed to provide a good barometer of one-off performances of participants in health education programmes. Still, further research would be needed to develop standards or criteria, which researchers could use to compare against their results. The idea here is that researchers would only need to, for example, observe one health education session to know from attention observations if the participant(s) are showing good attention levels. Hence, repeated one-off measurements of any of the output steps would give an indication of the performance of the participants.

This study also calls attention to the importance of some output steps. The study adapted the model to the setting of the study, and for that reason decided to measure skills acquisition in a more practical way, which is suggested by Rimer and Glassman "the evidence is that we should measure what people say and what they can do." Skills acquisition was considered the strongest and more important of all the output steps as suggested by John Elder, Professor of Health Promotion at San Diego State University Graduate School of Public (personal communication). Future studies could measure how strongly skills acquisition measured in this was a good predictor of behaviour change, in comparison to maybe recall or attitude change.

6.1.6 Summary of findings for this section:

- McGuire's model was found to be a useful and user-friendly conceptual framework for the process evaluation of the TSNI nutrition education program, suggesting that it would also be useful for any other nutrition education programs in developing countries, because it allowed for the breakdown of the communication persuasion process into steps which could be measured.
- The use of these output steps in turn, were able to detect differences and variability between/within the two intervention groups.
- The main limitation was the fact that output steps were not measured in the same subjects, which consequently meant that it was not possible and easy to explore the model in terms of the associations among the output steps, and with behaviour change.

6.2 The developed tools

A major contribution of this research study was the design and development of tools to measure the various output steps. A variety of tools were developed, ranging from simple questions to measurement scales and indexes, and systematic observations.

While conceptual definitions for each output step derive from literature review, the operational definitions were developed and adapted to the setting of this study. One of the reasons for this is that this is the first study to measure these output steps in a rural setting of a developing country and hence the majority of tools used in developed countries would not be appropriate for this study setting. Another reason is that the studies reviewed in chapter 4 for their methods to measure the same output steps, did not focus per se on tool development and their account of the methods was very brief.

This section looks at the strengths and weaknesses of each tool and suggests improvements.

6.2.1 Exposure/attendance measurement

Measuring attendance to classes was probably the most simple and straight forward measurement taken. Unlike the other methods/tools developed to measure the output steps, the method used to measure attendance levels was based entirely on other studies' [92, 93, 119]

There are nonetheless possible limitations with the way the measurement was carried out in the TSNi study:

Measurement is dependent on the health promoters and it is possible that they might have forgotten to measure attendance to some classes, or did not include mothers that arrived late to the class.

Attendance to group classes was considered positive only if the mother was exposed (i.e. was physically present on the class) to the various topics included in the nutrition education program curriculum. However, topics were sometimes repeated so that those mothers that did miss the first time could attend the repetition. However, attendance was measured as mothers' being exposed/ not exposed to a topic, regardless of repetition. So in fact, even though this research study reports a total of 12 group classes, with the repetitions there were probably near to 20 group classes. While this can be seen as a limitation of the study as it reduces the variability of the results,

this research study decided to follow McGuire's conceptual definition of exposure which implies solely exposure to the message: "was the mother exposed to the message?" If the "message" in question are the messages in the 12 topics then it does not matter if mothers missed the first class and only attended the repetition as they would be exposed to the message either way and what is considered important is for them to be exposed to that message. One could then have argued against the use of the term "attendance levels" instead of exposure. However, it was necessary for mothers to attend the group classes for them to be able to be exposed to the message(s)."

On the other hand this research study does not differentiate between those mothers that participated in two classes of the same topic (i.e. the first class and the repetition) and those that participated in just one, hence, it does not take into account repeated exposure. Perhaps future research could look at the proportion of attendance to the total number of topics (e.g. in the case of this study it would be 12 topics) and repetition. So for example, the comparison would be between subjects that have attended 6 out of 12 group classes and subjects that have also attended 6 group classes plus repetitions. However, the categorisation of these attendance levels with repetitions makes everything more complex and would imply an increase in the sample size.

In the same way, attendance to individual classes was considered positive only if the mother received the home visit from the health promoter. This implied that the previous home visits which the mother was found absent from her home were not accounted for because by the third visit she was found at home and the health promoter was able to deliver the message (or key messages in the case of IC). Thus again, it is the exposure to message that is important and not how many times it took the health promoter to visit the mother until she was able to deliver the class.

Moreover, by not measuring the actual attendance to group and individual classes regardless of exposure, the study is missing out in one factor: willingness of the mother. Following this thought, as was seen from the results, attendance levels to GC were significantly lower than attendance levels to IC. This is because attendance to GC was more determined by the mothers' willingness to receive the message since they had to make a much bigger effort to attend the GC, whilst attendance to IC was more determined by the health promoters visiting them and mothers did not need to make such a big effort.

6.2.2 Attention measurement

Probably one of the strengths in the measurement of this output step is that two methods, instead of just 1, were used to measure attention. Attention observations, the main method, were supplemented with attention recall questions. Most studies in health education only use one method. As suggested by Dr Casimir Ludwig (from the faculty of experimental psychology, University of Bristol): “monitoring gaze behaviour (referring to what it referred in this study as attention observation) seems like an a-priori valid and intuitive thing to do. It is possible that mothers may be making eye contact with the health promoter without paying any attention at all, but no single method is going to be watertight under the conditions of this study. It is also an added benefit to supplement it with additional measurements. (personal communication).” A study by Fischer *et al* (1989)[131] also used similar combined measures to examine whether subjects “read” and “recalled” warnings in tobacco advertising. They tested reading by recording eye tracking data during advertisement viewing, and recall by using standard post viewing recall techniques used in market research. Both measures were well-accepted market research techniques. The reason why Fischer *et al* (1989) used recall measures to supplement eye tracking is because it gives an extra measure for those cases where a low amount of learning is thought to have taken place. [131]

Attention observations in this research study were taken while the class was going on, giving an idea of what was happening through out the class, especially, if mothers were directing their attention to the messages. Attention recall, only tells us what the participant is able to retrieve after the class has finished, and hence is a step that comes after direction and maintenance of gaze. The two together make a more complete picture of the participants attention. In comparison to Fischer *et al* (1989)[131] study, this research study’s measure of visual attention was much more naturalistic, simpler, and applicable to the study setting. One could argue in favour of Fischer *et al* method as a more valid measure, however, their method was much more invasive which might have a higher negative effect on attention than a more naturalistic method like this study’s. The recall measures used in Fischer’s study and in this research study are different. The recall measure used in Fischer’s study consisted of specific questions regarding the content of the advertisements which is quite similar to this research study’s measure of recall. However, attention recall was much simpler and only consisted of asking mothers if they could remember if 2 certain words had been mentioned in that class.

Another strength is the use of observations to measure attention. Most of the studies in health education that have reported measuring attention, have done so by asking participants to recall features of the program (see section 4.3) or if they have measured

visual attention they have used sophisticated lab measures, like for example eye tracking devices. Only Peterson *et al* (1984)[135] and Palfrey *et al* (1981)[136] studies used observations similar to this research study. Classroom behaviour measurements, including attention, is also measured using systematic observations.[179] In a lab setting, recall is not even considered as a method to measure attention, one of the main methods consists of monitoring visual attention. Even though, the setting of this study was far from being a lab setting, it is still the most preferable method as gaze or sustained attention is the first sign of attention while recall, as mentioned by Dr Casimir Ludwig, involves a very strong memory component. Due to these observations it was possible to observe that mothers in a group setting acted differently than when in an one-to-one setting. During the pilot test it was observed that in an individual class mothers would mostly abstain from looking directly at the health promoter and would instead use additional behaviours more indicative of attention such as nodding and saying "yes". The study had to then, include these other signs in the attention observations. The same kind of behaviour was not observed during group classes. Hence, observations allowed the researcher to pick up different behaviours which would not happen with recall questions.

Regarding the application/administration of the observation tool in the field, no difficulties were encountered while making the observations. It was easy to observe mothers in the class without raising their suspicions and, because mothers were selected when they were already seated, only those who the researcher had a clear view, were selected. Mothers were only observed for 5 seconds. Peterson's study[135] also observed each participant for periods of 5 seconds. It is probably better to make observations that short. If observations for each mother were much longer and if during that observation the mother broke eye contact for an instant, it would make it more difficult to decide if that observation should be considered as "paid attention" or not. This researcher, hence found it better to do many quick observations instead of few longer ones. Recall questions were very simple and clear and mothers seemed to understand easily enough what was asked. It also made it easy that these recall questions had been standardised. The researcher made sure that the words asked for the bogus questions were known by the mothers. Hence, there were no cases where the mothers did not recognise the word. It was a good idea to add a bogus recall question as it avoids subjects getting points when they are only guessing the answer (i.e., saying yes to all the questions). Bogus question, like the non-bogus question was given a score of 1 if correct (i.e. if the mother answered that it was not mentioned in the class).

The main strength of both tools/methods, is that they were sensitive enough to measure variability within intervention groups. Moreover, attention levels measured with both methods were high. Results for Peterson's study also showed that attention levels measured with observations were very high.

One of the main weaknesses in the measure of attention is that even though attention recall was used to supplement attention observations, results for the two tools did not correlate with each other. Hence, instead of aggregating (i.e. adding up scores of the observations with the scores for attention recall) the results for both measures for each mother, they had to be looked separately. There are many reasons that could explain this lack of correlation. In cognitive science and neuroscience, attention can be reflected as two different modalities.[90] Perhaps, this is the case in this study, with observations measuring visual attention and recall measuring auditory attention. Or perhaps, the latter is more a reflection of memorisation, as discussed by Casimir Ludwig, which would mean that it is a step further than just visual attention. Another possibility is that while the observations were genuinely able to detect inattentiveness and attentiveness, there was response bias regarding recall questions (e.g. answering "yes" to all recall questions regardless of having heard the word said in the class or not). In that case, attention recall would not be reflecting attention levels and would not be a valid tool. However, a strong test-retest correlation coefficient for the recall tool proved the contrary. If mothers were answering at random, then the correlation between the two measurements would probably be low. A further possibility, is that attention recall is the tool that best reflects attention level and there was response bias during the observations in class (i.e. mothers were looking at the health promoter but not really paying attention). It is also possible that the number of observations and the number of recall questions asked per mother, were not enough to reflect attention levels. It is a bit more difficult to decide this with the observations, since already 10 observations were carried out. Both Peterson *et al* and Palfrey *et al* studies did not report the number of observations taken so its not possible to know what would be a sensible number. But only 2 recall questions were asked per mother, while the recall measure in other studies has consisted of more questions. Two questions reduces the variability within a group, and in fact, might produce a ceiling effect where the majority of scores are at or near the maximum possible for the test. In Fischer *et al* study, the two measures (i.e. recall and attention observations) had a statistically significant association with each other, however, the authors did not discuss the fact that the strength of the correlation, $r=0.252$, was low.

A recommendation for further research would be to increase observations to 20 per subject. This would also increase variability. As well, instead of only asking 2 recall questions, maybe 4 recall questions should be administered (i.e. two true and two bogus).

Another limitation of these two tools and the study design is that they were administered once for each mother. The limitation of this is that it only assesses the mother's performance in that class and it does not give an idea of her overall performance. What if the mother was more attentive on that day and not at all in the rest of the days? In fact, a test-retest reliability analysis of attentions observations in 10 mothers resulted in a very low coefficient, proving that attention levels vary a lot.

Hence, another recommendation for further research would be to carry out repetitions (at least one) of the observations and then average it per subject. The same with recall questions. They should be administered in another occasion on the same subject.

Attention-recall questions could also have been measured with block 3 output steps, and measured in the demonstration of the recipe. Measuring attention observation in a recipe demonstration would be very difficult and impractical as there is a lot of activity and mothers are moving around and hence it would be impossible to measure eye contacts.

As mentioned by Dr Casimir Ludwig, measuring eye contact seems like an a-priori valid method. Peterson *et al* study does not report any validation of the observation method. Palfrey *et al* was shown to have concurrent validity with the McCarthy Scales of Children's Abilities and predictive validity with regard to kindergarten Performance.[136] None, of the other studies that have also measured attention-recall reported measuring validity. Test-retest reliability was calculated for each tool. Results showed a high correlation coefficient for the attention-recall tool but a very low one for the attention observation tool. It was expected for attention observations to differ a lot between the two measures since so many other factors might affect each of the observations and the overall frame of mind of the participant in that day.

6.2.3 Comprehension measurement

One of the strengths in the measurement of this output step is again that two methods, instead of just 1, were used to measure it. These were based on Rimer and Glassman's study[78]. The advantage of using both measures is that they grasp different concepts. While comprehension-lexical meaning gives an idea on how much

of the key messages can the mother recall after just being exposed to the messages, comprehension-psychological meaning gives an idea of how much of more complex concepts can they recall after 2-4 months. Also, while comprehension-lexical meaning was a reflection of the didactic classes, the comprehension-psychological meaning was more a reflection of the practical classes. Because of these differences in both measures, unlike the attention output step, the measures of comprehension were not developed to supplement each other. This was also the case in Rimer and Glassman study.

Like Rimer and Glassman's study[78], the comprehension psychological meaning tool consisted of a composite of several questions, an index. Questions were also of the same type as Rimer and Glassman study's. Like Rimer and Glassman's study, each question was given a score of 1 if answered correctly. However, unlike Rimer and Glassman's study, because the questions were linked to the recipe they carried out for the task, mothers received different numbers of questions. Even though a proportion was calculated for each mother resulting in a comparable figure for all women, it does not take into account the fact that perhaps the effort of answering 8 questions is more than only answering 5. Perhaps, one way of avoiding this would have been to ask the same questions regardless of the number of ingredients in the recipe, for example, just ask questions regarding the first two ingredients in a recipe with 4 ingredients which should match the questions for those recipes with only two ingredients.

There did not seem to be any difficulty in the administration of the comprehension-psychological meaning questions as they had been previously tested for their clarity and cultural context. Answers (including the wrong ones) also proved that the mothers did understand the question since they were within the type of answers expected. However, mothers seemed to have more problems in answering some items than others. The items addressing concepts on the benefits of certain food groups for the child's health were the most difficult to answer. Probably because they required the mothers to give very precise answers. So for example, when asked why is the carbohydrate rich food group good for the child, the answer had to be that it gave them strength. This is the message that was constantly given to them, especially in IC. However, if they answered that it gave them energy, even though this may be true, it was not the answer expected. Ashworth and Feachem (1985)[22] also argued the same point: "education about different nutrients and their functions, and about the principles of a balanced diet, is not readily understood by most audiences in developing countries." [22] Mothers found it easier to answer the items addressing their ability of correctly recalling a substitute for the fat-rich (coconut milk or peanut, depending on the

recipe) and protein-rich (fish) ingredients used in the recipe. However, when asked to give a substitute for the fish, they mostly chose incomplete proteins-rich foods which were also fat rich, like for example, cashew nut and seeds, and not for example meat or beans. Sometimes they mentioned egg. Perhaps this is because animal sources are not readily available, while cashew nuts and seeds are. Mothers also seemed to find it difficult to remember which of the three recipes taught initially was the best for the child, but even more difficult, was to explain why was it the best. The researcher believes that perhaps the first set of items were not necessary and relevant. Maybe it is not important if mothers mistakenly said that carbohydrates rich foods gave energy to the child instead of strength because they could not remember the exact message, as long as they know that it is important for the child to eat them everyday. Perhaps, it was more important to verify if mothers understood the reasons behind each of the cooking steps (i.e. the ones assessed for skills acquisition). There is a possibility that if the mother does not understand that the reason they were taught to boil the orange-fleshed sweet potato (OFSP) with their skin is to avoid the leaching of beta-carotene (referred to as vitamin A in the TSNI study) than they might not do it in the future. Instead, questions should have been related to each of the guidelines for the skills acquisition task.

Results from this study suggest that both comprehension tools were sensitive enough to detect variability within the intervention groups.

Understanding levels for comprehension lexical meaning were in general much higher than comprehension –psychological meaning. It would be expected for proportions of correct answers to the comprehension lexical meaning measurement to be higher since questions contained much less complex concepts than the comprehension-psychological meaning questions. Moreover, whilst comprehension lexical meaning was administered right after the mothers were exposed to the messages, comprehension psychological meaning was only administered 2-4 months after the mothers had been exposed to the messages, which probably means that mothers might have forgotten many of the concepts taught to them. We do not know however, what made the answering of the comprehension-psychological meaning questions more difficult, the fact that they contained complex concepts or that they were asked 2-4 months after the exposure to the messages. Perhaps, the same comprehension-lexical meaning questions could have been asked 2-4 months after exposure in order to measure ability to recall these more easy questions with the time factor.

It is possible that while the lack of a significant difference between the two interventions for the comprehension psychological meaning was not due to the tool itself, the same did not happen with comprehension lexical meaning. The lack of a significant difference between the two interventions for the comprehension lexical meaning could have derived from weaknesses in the tool.

This weakness comes from the fact that even though the comprehension lexical meaning tool consisted of 3 questions, only one question was administered to the mothers. This implies limited variability of answers and once again, a ceiling effect which might mean that the question on the test was insufficiently difficult to measure true ability or comprehension. Other studies that have also attempted to measure comprehension lexical meaning have asked more than just 1 question. Rimer and Glassman's study measured comprehension lexical meaning by administering 5 questions and then scoring them. Moreover, in this research study, the question was administered only once and might reflect mothers disposition or attention to that class only.

A recommendation would be to make the comprehension lexical meaning tool into an index of several questions (maybe 3), repeat this measurement in another occasion (but not necessarily the same questions) and calculate an average for each mother.

All comprehension lexical meaning questions asked were clear, simple enough and sensible for the setting of the study as had been previously tested. Even though it was open ended questions, certain answers were expected. For example, for the question on "how many main meals should child x have per day", correct answers could be: 3; should eat in the morning, afternoon and dinner; should have breakfast, lunch and dinner; or even (as some mothers said) should have 3 main meals and snack in between them. Like the comprehension psychological meaning questions, results show that there is also a difference in the difficulty of the comprehension lexical meaning questions, with lower proportion of correct answers to the question "give me 3 examples of vitamin A-rich foods. A possible explanation for this is that there were more items to recall for this question. But also the key message on vitamin A rich foods was the first one given in the class, meaning that it is possible that by the end of the class mothers mixed the key messages.

Perhaps, one limitation regarding the answers to the comprehension lexical meaning questions is that one cannot be sure if the mothers really understood the key message or were only retrieving memorised information. One could argue that the reason why

the proportion of mothers correctly answering the question was high was because the questions were too easy. However, all questions asked were related to one of the key messages of the topics taught which were designed by the government. Moreover, the 3 questions varied in their level of difficulty rendering the tool able to give an idea of the level of answers. We know that if mothers had to recall more than 1 item, they found it more difficult.

6.2.4 Recall and skills acquisition measurement

McGuire, never defined the ways in which each output step should be used. He mainly devised the sequence of output steps which should be used to evaluate health communication programs. Many of the studies that have used his model, specifically the output steps, have adapted it to their objectives. Russon and Koehly study (1995)[85] for example created their own operational definitions for each output step. They also put together some of the output steps. For example, they measured “yielding” which included acquiring relevant skills and attitudes. This research study also developed a task to measure comprehension-psychological meaning, skills acquisition and recall. Even though these output steps measured different concepts, they were all connected by the same task.

It has been shown by many that many child feeding practices in developing countries are considered inappropriate[22, 49]. The TSNi baseline study confirmed that this was also the case in the TSNi study areas.[16] These inappropriate dietary practices could derive from lack of knowledge or lack of resources.⁸ Taking the first into consideration, recipe demonstration which was part of the TSNi nutrition education program curriculum and this research study's assessment task is a way of dealing with the lack of knowledge. Moreover, when introducing a new cultivar such as the OFSP, which unlike the white-fleshed sweet potato, is rich in beta-carotene, care needs to be taken when cooking it in order to retain its vitamin A content. The recipe task used in this research study assessed mothers ability to recall the recipes, their ingredients, the various steps to carry out the recipe and actually performing it. None of the studies that included recipe demonstration or food preparation skills in developing countries evaluated using a similar task. Asking subjects to perform a behaviour taught to them is the only way to make sure that they are capable of doing it. Moreover, the fact that they have to perform it and not just hear about it might ensure that they remember it. As shown by Nagarajan *et al* (1989)[150], which compared glucose concentrations of a oral rehydration solution (ORS) carried out by 3 different groups, the group of mothers that received live demonstrations prepared ORS with a glucose concentration closer to

the ideal concentration in comparison to the groups that only received verbal instructions. However, Jacoby *et al* (1994)[180] study results showed that the addition of cooking demonstration to a simple oral instruction did not increase the efficacy of the intervention. However, in Jacoby study mothers were not asked to repeat the recipes on their own. Hence, the development of this research study's task and its evaluation using the 3 output steps is an added strength. In developed countries some studies have shown that hands on work with food is more linked with behavioural change. [181, 182]

One of the content limitations of the tools to evaluate this task, already mentioned above, is the lack of questions regarding the understanding of the reasons behind the steps taken for the preparation of the recipes.

All three output steps measured for this task consisted of a composite of several questions (i.e. an index) or, in the case of skills acquisition, guidelines that they had to follow. This is the case with all other studies that have also measured these same output steps.

The output step recall in this study is differentiated from attention-recall and comprehension-lexical meaning by the time factor. While the other two were more a measure of short-term recall, recall was a measure of longer-term recall. Nonetheless, comprehension-psychological meaning could also have been used as a measure of long-term recall, which was the case of Rimer and Glassman's study. The recall tool however, was composed of relatively easy items. All items in the recall tool were related to the recipes. To avoid ambiguity in the answers which could interfere with the allocation of scores, rules were created for when answers should receive scores or not. For example, if mothers failed to recall one of the ingredients used in the chosen recipe, then they would not receive any score.

The administration of the recall tool included the masking of the 4 food groups' poster and asking mothers to identify to which of the masked food groups did the recipe ingredients belong to. This was also a technique used in Fischer *et al* study where they masked the warnings in the advertisement for the recognition test. This is what is referred to as cued recall. Since mothers in this research study did not seem to do so well when asked about the benefits of the food groups for the child's health, it was thought that maybe it would be better to have them point out to which food group did each of the ingredients belong to than making them have to answer that it belonged to that food group with benefit X. At least it gives an idea to the mothers' knowledge of

which foods belongs to which food groups. 73.8% of the mothers (out of 80) were able to point out that OFSP belonged to the right (masked) food group, and 60% were able to point out the food group to which the second ingredient belonged to. Mothers in this research study did much better than mothers in Andrien study[183], where 86% were unable to classify foods into 3 food different groups.[183] Nonetheless, the masked technique used in this research study could be modified, and instead of covering completely the food groups in the poster, the mothers should have been allowed to see the poster with the food groups containing all the food, except those used in the recipe. In this way, we are testing if she remembers with which other foods are the foods used in the recipe associated with. This means, that if mothers puts fish (one of the foods used in one of the recipes) in the right food group, she also knows that meat, eggs, beans are also possible substitutes.

Of interest are the different results between the recall output step and the comprehension-lexical meaning. One would have expected the proportion of correct answers to the comprehension-lexical meaning question "give me 3 examples of vitamin A-rich foods" to be higher than the recall question "what were the ingredients used in the recipe chosen today?" simply because of the time factor. However, only 50.6% of the mothers (40 out of 79 mothers, regardless of intervention group) got the correct answer for the comprehension-lexical meaning question. While with the recall question, 93.8% of the mothers (46.9 out of 80 mothers, regardless of intervention group) managed to correctly recall all of the ingredients used in the chosen recipe for the task. Even for the easiest comprehension question "how many main meals should child x have per day?" Around seventy nine percent (52 out of 66 mothers) got the correct answer, which is still less than the recall question. Moreover regarding the other recall questions, 83.8% of the mothers (41.9 out of 80 mothers) correctly recalled that 3 recipes were carried out in the demonstration, and between 75-80% of the mothers correctly recalled all the ingredients in the recipes carried out in the demonstration. Obviously time factor was not so strong to deter answers for the recall questions, or there was another factor that had a bigger weight on the answers to the recall question. The only other difference between the recall tool and the comprehension-lexical meaning tool, is that the recall questions were associated with an exercise (i.e. recipe demonstrations) carried out by the mothers. So, instead of just listening to the health promoters explaining how to carry out a certain recipe with certain ingredients, mothers had hands-on experience at making them. Thus, perhaps, it is this hands-one experience that helps mother to better recall the information. If this is the case, it confirms again the importance to carry out this kind of task (i.e. demonstrating how to prepare weaning recipes and then ask mothers to prepare them themselves).

Like comprehension-psychological meaning, because some of the recall questions were related to the chosen task recipe mothers received different number of questions. Even though a proportion was calculated for each mother which adjusted for the different number of questions, it does not take into account the fact that perhaps the effort of answering 9 questions is more than only answering 7. Perhaps, one way of avoiding this would have been to ask the same questions regardless of the number of ingredients in the recipe, for example, just ask questions regarding the first two ingredients in a recipe with 4 ingredients which should match the questions for those recipes with only two ingredients. That would ensure that all the mothers were asked similar questions.

The skills acquisition tool consisted of 8 steps (regardless of cooking method) which the mother had to follow. However, only 6 of the steps were scored as the other two were irrelevant for some recipes. Unlike the recall and comprehension-psychological meaning tools, all mothers had to follow the same steps. All 6 steps for both cooking methods were related to the conservation of beta-carotene in the porridge. The tool was not difficult to apply, the researcher just had to be careful to not say anything that might make it easier for some mothers. For that reason, a plan was designed before approaching the mothers and this was shared with the health promoters. This plan consisted of making sure that the mothers themselves chose the recipe that they wanted to prepare and recalled which ingredients were part of that recipe. She did not receive any help from the researcher or the health promoter. As well, during the preparation of the recipe, all mothers were refused any kind of help. A pilot with two women was carried out to test the flow of the task, which consisted of approaching the mother to ask her to chose a recipe and the actual preparation of the recipe. The pilot also allowed the researcher to decide when to decide that the mother had performed well in that step or not. Unlike the other two tools measured with this task, the skills acquisition score also included half points. This was the case for 3 out of the 6 steps for the normal boiling recipes and. For example, the mother would get a score if she added enough water to only cover half of the potatoes, but if she added more than that but still it had not covered the potatoes completely, then she would get $\frac{1}{2}$ point. This study tried to control for ambiguity by having only one person assessing mothers' performance in each step and by having a standardised rule of what should be considered as full point, $\frac{1}{2}$ point or no points. There could have been a problem in assessing two cooking methods as one could have been easier than the other. however, there was no statistically significant difference in the scores between the two cooking methods.

Skill acquisition scores were higher than was expected. In fact, the proportion of mothers that got full scores (i.e. 6) for skills acquisition was higher than the proportion of mothers that got full scores for the recall output step. Mothers seemed to remember the appropriate cooking methods for each of the recipes, since only 1 mother used the wrong cooking method. Moreover, 69% of those using the steaming method, and 60% of those using the boiling method were able to add the right quantity of water. Also, 65.5% of those using the steaming method, and 62% of those using the boiling method were able to identify the right cooking time for the orange-fleshed sweet potato (i.e. had to be a bit hard).

One possible argument for the unexpected high scores for the skills acquisition tool is that maybe mothers had already prepared these recipes and hence it was not something new for them, and because there is not a control group, it is not possible to know if it was really the nutrition education program that improved these skills. However, the orange-fleshed sweet potato (OFSP) was only introduced with the TSNI study, and all the steps taught for the recipes were related to the conservation of beta-carotene in the OFSP. Moreover, TSNI baseline data showed that weaning porridges consisted mainly of cassava flour, or another flour, hence, there was no need to follow certain steps to preserve as much as possible the beta-carotene content of the porridge. But the most important fact that refutes this argument, is that there was a very strong difference in the skills scores between the two intervention groups.

Results from this study suggest that both recall and skills acquisition tools were sensitive enough to detect variability within the intervention groups.

6.2.5 Attitude measurement

While most of the other tools were an adaptation of tools used in other studies, the attitude tool developed in this study is probably the only one that was based entirely on a reference[116] which described in detail how to construct attitude scales.

There have not been any up to date studies that have measured attitudes towards breast-feeding and complementary feeding in Mozambique or even neighbouring countries. For this reason, all items had to be devised *de novo* for this tool. Nonetheless, the majority of items were related to the nutrition education teachings.

An initial pilot study of the attitude tool was carried out, but the questionnaire proved satisfactory and remained unchanged for the final survey except for the omission of two

items which were irrelevant since they were to be used as control items for section A of the tool. A test-retest reliability measure was also carried out which showed that scores were stable over time (i.e. a period of a week). The internal consistency of the tool was also measured for each section of the tool. However, this was only measured after the tool had been administered to the study group. The usual procedure is to measure the internal consistency after the pilot study in order to decide which items should be excluded. Nonetheless, internal consistency was high among the items in section A and borderline among the items in section B.

The attitude tool developed for this study had many strengths. First, instead of just asking mothers for their attitudes towards an item, they were also asked to give their reasons behind their attitudes. This allowed the researcher to verify answers according to the question and confirm if mothers did really understand the question. Second, the tool included story telling and vignettes which disguised the purpose of the tool and also facilitated the understanding of each item. Third, bogus items (including an item where high agreement was expected) were added to control for mothers answering without really thinking about it. They were very useful at detecting if the mothers really agreed with all the items or if there was response bias (i.e. mother not giving the real answer). The cases where this happened were excluded from data analysis. Fourth, items were negatively-keyed in order to avoid socially desirable answers.

It was also a good idea to use the 5-point Likert scale. The fact that respondents used the full range of values and not only the extreme points (i.e. 1 and 5) suggests that they weighed their answers, further suggesting that the 5-point Likert scale was adequate for the setting of this study.

The attitude tool was easy to administer, and the pilot test showed that items were culturally acceptable, even though those in section A involved very traditional and some taboo practices.

At least 2 items were not related to the nutrition education program, and maybe were redundant in this tool. The reason they were added to the tool was to give a general idea of the importance that the mothers placed on breastfeeding and complementary feeding over other activities (i.e. going to the field) and attitudes (i.e. that the father has to be better fed than the child), however, the researcher did not know the baseline attitudes for these two items. There is the danger of these items lowering the overall score for each section and hence, making the difference between the intervention groups and the control group statistically non significant.

Possibly there is one item that might have been unclear to the mothers since result was very unexpected. Mothers in both intervention groups were exposed to the message that they should add a source of fat to the child's 3 main meals every day. When asked in the attitude tool if they agreed that it was enough for the child to eat 3 times a week a source of fat, a very high proportion of mothers in both interventions (82.7% in intervention I and 83.8% in intervention II) agreed while a significant smaller proportion (70%) in the control group agreed to it. It is possible that intervention mothers got confused with messages using both 3 times a day and 3 times week.

The main finding is that the attitude tool was sensitive enough to detect variability within the 3 treatment groups.

6.2.6 Measuring reliability and validity

A major limitation in the development of tools in the current study is that it did not ascertain the validity and reliability of a lot of the tools developed for the study.

Reliability measures were only carried out for some of the output steps: test-retest for attention observation, attention recall and the attitude scale; and, internal consistency was carried out for the attitude scale. This means that reliability was not measured for comprehension lexical meaning, comprehension psychological meaning, recall and skills acquisition. A test-retest measurement should have been carried out for these tools to see at least, if results could be reproduced in the same mothers. The main reason why this was not done so for comprehension-psychological meaning, recall and skills acquisition was due to time and resource constrain, but a bigger effort should have been taken to carry out these measurements.

None of the tools were validated. All of the tools were developed *de novo* for the current research and for that reason it was difficult to validate it.

It could be argued, and it was mentioned by Dr Casimir Ludwig, that it is possible that mothers could have been making eye contact with the health promoter without paying any attention at all. A way of validating it would be to compare it against one of the more used lab based techniques, such as an electroencephalogram which would record event-related potentials and reaction times. However, this would be extremely invasive for this study's participants, and again, one would not know to what extent is this more invasive technique having an effect on attention. So, as it is, there is no

possible way of verifying Ludwig's statement. However, as mentioned by Dr Casimir Ludwig, measuring eye contact seems like an a-priori valid method.

It could also be argued that both attention observation and attention recall should be able to validate against each other, as a person that was indeed making eye contact and hence, attentive, should be able to recall correctly if a word or a message was said correctly. However, as mentioned before number of observations and recall questions might not have been enough to ascertain this theory and hence, future research taking this into account could check one method against the other in order to measure their validity.

Face validity could have been measured for comprehension lexical meaning, comprehension-psychological meaning, recall and skills acquisition, whereby a panel of experts from the London School of Hygiene and Tropical Medicine could have reviewed the tools after they had been developed and make judgements about the extent to which a particular tool measured what it was intended to measure.

Construct validity could have been measured for the attitude scale, with the scale being administered to a population where attitudes were already thought to be positive (i.e. health promoters). Concurrent validity could also have been measured for this instrument, with attitudes compared against some behaviours.

6.2.7 Summary of findings for this section

- Developed series of tools which were able to detect variability within intervention groups;
- All of these tools could be adapted for use in other nutrition education projects in developing countries;
- All tools were easy to administer and were culturally acceptable;
- Future research should include more observations for attention, more attention recall and comprehension lexical meaning questions;
- Future research should include more than one administration of attention observations, attention recall, and comprehension lexical meaning;
- The recipe task used to measure comprehension-psychological meaning, recall and skills acquisition should be included in future nutrition education programs which focus on complementary feeding since it has shown to improve recall abilities of mothers. It also gives us a more complete idea of, not only what the participant has learned, but what he/she is able to do;

- Cooking classes alone along with the explanation of the various food groups and their importance would be enough if there was not the need to change certain beliefs which hinders behaviour change;
- Future research should validate and measure the reliability of each tool since majority of studies reviewed for this thesis which have also attempted to measure the same output steps have not done so.

6.3 Which intervention group is more effective?

The term effectiveness in this case is very subjective to what the researcher is trying to measure. In the case of this research study, the researcher carried out a process evaluation of the nutrition education program delivered with intervention I and with intervention II. The process evaluation specifically measured McGuire's output steps in each of the intervention groups. Effectiveness in this study can then be measured as the performance of mothers in each of the output steps which is then compared between the two intervention groups.

The important question to ask is if performance in the group that only received group classes (i.e. intervention I) is as good as the group that received group classes plus individual classes (i.e. intervention II). This research study's hypothesis is that performance of intervention I mothers would be significantly lower for all output steps when compared with intervention II mothers', simply because intervention II mothers had the added benefit of receiving the individual classes apart from the group classes. This section will discuss the results which support or contradict this hypothesis.

A review of the literature in section 2.3 of this thesis did not retrieve any other studies that have also attempted to compare a group class intervention with a group class plus individual class intervention for any of the output steps. Hence, this is the first study of the kind.

6.3.1 Exposure/attendance rates

- *Do attendance rates differ significantly between the approach-intervention groups (i.e. GC intervention I, GC intervention II and IC) and between the intervention groups?*

Supporting the hypothesis, attendance rates to individual classes were significantly higher than attendance rates to intervention I group classes. This was also supported by two other studies that compared completion of sessions and adherence to a nutrition education program in the individual setting against a group one. [92, 93] A possible explanation for why this is the case is that while for the group class approach the mothers were warned about the class and had to turn up for the class which meant that attendance depended on mothers' willingness to attend, for the individual class approach the mothers were also warned about the home visit, but then it was the health promoter that would visit the mother. Hence, the TSNI study had a much bigger control over attendance to IC. This is also discussed in Bull *et al* (2001) study: "in many instances message exposure, the first step, is under the control of the researcher, not the participants." [87]

Attendance rates to group classes by intervention II mothers were significantly lower than attendance rates to group classes by intervention I mothers. This suggests that attending IC could have impeded the attendance to GC. A possible reason for this is that mothers in rural areas of Mozambique have not only to care for their children, husband and home in general, but are also actively engaged in agriculture. Hence, with these two main important roles to fulfil, there is hardly any time for them to attend individual classes, let alone to attend the group classes as well. When those intervention II mothers that attended less than 6 group classes were asked for the reasons that led them to miss the classes, 33.7% said they were busy, 28.5% said they were ill and 22.5% said they had gone away from the village.

It has been reiterated, and described by McGuire himself, that for a health communication to even begin having an impact, subjects need to be exposed to the messages. However, if attendance to one approach (in this case GC) might be affected by attendance to the other approach (in this case IC), then, this raises many questions on whether it is worth having a combined intervention.

However, when results for GC intervention II and IC were combined and compared with intervention I, the proportion of mothers in **intervention II** attending classes was significantly higher than the proportion of mothers in **intervention I**. Hence, even though mothers in intervention II had to attend both GC and IC, their overall attendance rates to the nutrition program was significantly higher than Intervention I mothers.

Another important issue to discuss is that both intervention I and II attendance level to the first group class topic was relatively high with around 80% of mothers (88% for

intervention I) attending it. However, at the end of the program, attendance level decreased, and around 60% of mothers (65% for intervention II group) attended the last topic. This was also supported by another study that found adherence in the first year of the program to be much higher than in the second year.[93] Whilst there was a drop for both intervention groups, this was only statistically significant for intervention I. Hence, even though intervention II mothers' attendance to GC was lower than intervention I, it seems that they were more motivated to keep on attending them. A possible explanation for the drop in attendance levels for both intervention groups could be due to an increased lack of interest, especially for intervention I mothers as 10.5% in intervention I admitted lack of interest as one of the reasons for not attending group classes against only 2% for intervention II mothers. Another possible explanation, is that even though attending IC impeded attendance to GC, the inverse relationship happened for those mothers with high attendance to IC, as results have also shown a strong positive correlation between attendance to IC and attendance to GC suggesting that those intervention II mothers with high IC attendance were motivated to attend GC.

The most probable explanation for the drop in attendance rates to group classes for both intervention groups was engagement in agricultural activities during the rainy season. After looking at figure 5.1 (results chapter 5) it is possible to see that the peak in the drop coincides, for both intervention groups, with the rainy season. Topics 5, 6, 7 were given approximately around the rainy season, which is time that women dedicate their time exclusively to planting and weeding.

6.3.2 Block 2 output steps

- *Do attention scores differ significantly between the approach-intervention groups and intervention groups?*
- *Do understanding scores differ significantly between the approach-intervention groups and intervention groups?*

Attention observation scores

Before adjusting for confounding (i.e. class duration) attention levels in IC were significantly higher than attention levels in GC intervention I. After adjusting for class duration, attention levels in IC were only marginally higher than attention levels in GC intervention I. However, attention levels before and after adjusting for confounding were not different between GC intervention II and GC intervention I, but most importantly, not different between the **intervention groups**.

Hence, results did not support the hypothesis that attention scores would be higher for intervention II. We would have expected mothers receiving the individual classes to show higher attention levels, in view of the fact that in an one-to-one setting the health promoter is 100% dedicated to the mother and also, there is relatively less disturbance. This was also suggested by Blatchford (2003) [179] study which included systematic observations of student's behaviour in large classes (i.e. 20-30 students) and also in much smaller classes (i.e. 4-5 students) and showed that in the larger classes there was more pupil inattentiveness and off-task behaviour than in the smaller classes.[179]

A possible explanation for the lack of difference in attention scores between the intervention groups comes from the fact that sample size calculations were initially calculated to detect a difference between the two approaches (GC intervention I and II pooled together versus IC), and when this comparison was carried out, attention scores in IC were significantly higher than attention scores in the GC approach. However, because the comparison is in fact between intervention groups, the GC approach sample size was sub divided into GC intervention I and II. A sample size of 91 which was expected to detect a difference of 20%, was broken down and only 42 mothers belonged to GC intervention I, much smaller than the number retrieved from the sample size calculations.

This suggests that: either attention scores in intervention II were not really significantly higher than attention scores in intervention I or the difference (of 11%) that the study detected could have been significant if the sample size for intervention I was bigger.

Results in this research study have also shown that duration of class seemed to have a negative association with attention scores. However, while duration of class seemed to have a detrimental effect on attention scores measured in GC intervention I and II, it did not seem to affect IC. Cole *et al* (1980)[134] and McDonald *et al* (2004)[184] showed respectively that mothers accompanied by a baby or small child may reduce concentration or might keep mothers from looking at a display, and that noise and interruption hinder learning in health education sessions. This research study also looked at similar factors. Noise/disturbance deriving from women standing up and babies crying during group classes did not seem to affect attention scores, neither size of group classes. However, the measurements of any of these 4 possible confounders were not carried out systematically and because of missing values, especially with duration of class in GC-intervention I where only 22 mothers (out of 42) were measured, statistical significance might be affected. Moreover, results might not be representative of the study sample population.

Socio-demographic characteristics did not seem to be associated with the outcome, but only a few socio-demographic factors, from many collected by the TSNI study, were explored, and therefore it is not possible to know if other socio-demographic variables could have had an effect.

Attention-recall

Before adjusting for confounding (i.e. class duration) attention recall in IC was significantly higher (by 20%) than attention recall in GC intervention I. As argued by McGuire, the repetition of messages, which was one of the qualities of the IC approach, aids memorisation and recall.[71] After adjusting for class duration attention recall in IC was only marginally higher than attention recall in GC intervention I. Attention recall before and after adjusting for confounding was not different between GC intervention II and GC intervention I. Most importantly, before adjusting for confounding attention recall in **intervention II** was significantly higher than attention recall in **intervention I**. But after adjusting for confounding attention recall was not different between the **intervention groups**.

Hence, results did not support the hypothesis that attention recall would be higher for intervention II. However, care must be taken when interpreting the results after adjusting for confounding as they were run on a very small sample size (especially for GC intervention I) due to missing values for class duration.

There are, therefore, 3 possible explanations: i) attention recall in intervention II was truly higher than attention recall in intervention I because confounding tests were only carried out on a smaller sample size and are possibly not representative; or, ii) class duration was really confounding the association between intervention groups and thus in fact, there is no difference between the intervention groups; or, iii) even with class duration as a true confounder, the difference between the two intervention groups could have been significant if the sample size for GC intervention I was bigger.

Results have also shown that duration of class only seemed to have a negative association with attention recall of GC intervention II mothers. It did not seem to affect IC and intervention I mothers. Noise/disturbance deriving from women standing up and babies crying during group classes did not seem to affect attention-recall, nor did size of group classes. However, the measurements of these possible confounders were not carried out systematically and because of missing values, especially with duration of

class in GC-intervention I where only 22 mothers (out of 42) were measured, it is difficult to say if the results are representative of the study sample population.

Comprehension – lexical meaning

Being able to recall key messages goes a step further than paying attention to messages as the former pre-supposes that mothers have paid attention.

Before adjusting for question type the proportion of correct answers in IC was almost 16% higher than in GC intervention I, however, this difference was not statistically significant. It would be expected for IC approach to do better than the GC approach since as discussed by McGuire, repetition of messages aids their recall.[71] There was also no statistically significant difference between GC intervention I and GC intervention II and, most importantly, no statistically significant difference between the two **Intervention groups**.

After adjusting for question type, the proportion of correct answers in IC was significantly higher than for GC intervention I. A possible reason for this is that while many more messages were given in GC and not so much in an orderly fashion, IC was based solely in giving 12 recommendations, which were in fact key messages, given in an orderly fashion. Hence, it is possible that it was easier for mothers to pick up key messages from the IC approach. This is also suggested by Naveie's (1994)[84] study which showed that key messages were better retained and the more messages there is in a topic, the more difficult it is to retain the messages.[84]

Additionally, after adjusting for question type, the proportion of correct answers in GC intervention II was also significantly higher than for GC intervention I which suggests that there was an added benefit of receiving IC as intervention II mothers' ability to recall key messages in GC also significantly improved.

Most importantly, after adjusting for question type, the proportion of correct answers in **intervention II** was significantly higher than for **intervention I**.

The proportion of correct answers to the comprehension questions varied according to the number of items that needed to be recalled in each question. Mothers succeeded in recalling the one-item questions (such as "how many main meals should a child of this age have?"), but when asked to recall 3 items (like for example when they were asked to recall 3 vitamin-A rich foods), the proportion of correct answers decreased.

Class duration, and class disturbances (i.e. women standing up and children crying) also seem have a synergistic effect. After adjusting for these, the proportion of correct answers in GC intervention II was significantly higher than for GC intervention I. However, the measurements of any of these 3 possible confounders were not carried out systematically and because of missing values, especially with duration of class in GC-intervention I where only 22 mothers (out of 42) were measured, it is difficult to say if the results are representative of the study sample population.

Block 2 output steps - Limitations and suggestions for future research:

The measurement of these 3 output steps, also referred to as block 2-output steps, which were measured together, seemed to have many methodological and study design limitations and probably from all the output steps are the ones with more limitations. The methodological limitations were already discussed in the previous sections. The main study design limitation is the small sample size for intervention I which might have led to results not having enough power. Secondly, some exposure variables, specifically class duration and disturbance levels were not collected in a systematic way and there were a lot of missing values which pushed down even more the sample size. Hence, the main suggestion for future research would be to look at the differences between these two interventions but with a bigger sample size for both GC intervention I and GC intervention II. Moreover, care should be taken on the collection of possible confounders such as class duration, and disturbance levels. Since there is now information on effect size and variance it is possible to calculate sample sizes correctly for any future study.

Another limitation in the study design was the absence of information (i.e. study identification) for those mothers measured in group classes. This limitation has for example made it impossible to explore the associations between attention and comprehension-lexical meaning with socio-demographic variables such as age, education level, and household income, and to assess whether selected mothers were representative of the whole study population. Again, future research should collect this information on the study subjects so that results can be associated and adjusted for these exposure factors.

6.3.3 Block 3 output steps

- *Do comprehension-psychological meaning scores differ significantly between the intervention groups?*
- *Do skills acquisition scores differ significantly between the intervention groups?*
- *Do recall scores differ significantly between the intervention groups?*

Whilst with the block 2 output steps and exposure it was possible to breakdown the intervention groups into their respective approaches (i.e. GC intervention I, GC intervention II and IC), with block 3 output steps (comprehension-psychological meaning, skills acquisition and recall) it is only possible to discuss results at the intervention group level, hence, intervention I versus intervention II. This is because intervention II mothers received the initial demonstration in an one-to-one approach and not in a group class approach. So, the GC intervention II group was not measured in block 3 output steps. What is being compared in block 3 output steps are the different effects of receiving the recipe task on a group setting against receiving it in an one-to-one setting. Hence, it will not be possible to extrapolate, which was possible in the previous output step measurements, how differently the mothers in intervention II would have performed from intervention I if they had received the recipe demonstration task in a group setting.

Comprehension-psychological meaning

Being able to process the information, thinking on how to apply it and how relevant it is, goes a step further than only being able to recall (key) messages.

Before adjusting for father's education level, **intervention II** scores were 10% higher than **intervention I**, but this difference was not statistically significant. After adjusting for fathers education level, scores for **intervention II** were 16% but not statistically significantly higher than **intervention I**.

After adjusting for time interval (the time that passed between the day that mothers received the demonstrations to the day that they had to carry out the recipe task on their own) scores for intervention II decreased.

After adjusting for both confounders, scores for intervention II were only approximately 12% higher than intervention I. But this difference is not of statistical significance. This suggests that there is really no difference between the two intervention groups, but maybe there was not enough power to detect a difference.

Comprehension-psychological meaning tool was composed of several questions, but only the question *which recipe (of those demonstrated) was the best for the child and why?*, produced significantly higher answers for intervention II. All the other questions produced similar responses between the two intervention groups. Possibly, the one-to-one setting facilitated the comprehension of some messages. However, it was expected that intervention II would also score significantly higher for the previous questions as these were messages given repeatedly in the individual classes.

Results suggest that mothers had more problems in answering comprehension-psychological meaning questions than comprehension-lexical meaning questions. This reflects the complexity and difficulty of the comprehension-psychological questions. This was also suggested by Navaie's (1994)[84] study which showed that simple messages without technical terms were better comprehended (i.e. less difficulty in recalling the messages and managed to retain the content) than more complex messages containing technical terms. But comprehension-psychological meaning also has a memory component as the tool was administered 2-4 months after mothers were exposed to the message.

Recall

Being able to recall messages 2-4 months after exposure goes a step further than only being able to recall messages after immediate exposure.

Before adjusting for father's education level intervention II scores were approximately 10% higher than intervention I, but this difference was not statistically significant. After adjusting for fathers education level the difference in scores did not change much (11% higher for intervention II).

After adjusting for time interval scores for intervention II decreased. After adjusting for both confounders, scores for intervention II were only 9% higher than intervention I. This suggests that there is really no difference between the two intervention groups, but maybe there was not enough power to detect a difference.

Recall questions were probably easy enough that intervention I was able to answer them as well as intervention II. The supposed added benefit of receiving the demonstration in a one-to-one setting did not make a significant difference to recall scores.

There was a much higher proportion of mothers answering correctly the recall questions in comparison to comprehension-psychological meaning. This suggests, that comprehension-psychological meaning questions were more difficult than the recall questions. The reason for this difference is that recall questions were much simpler than comprehension-psychological questions, and also, recall was actually cued recall, whereby a cue is used to stimulate recall. The fact that recall scores were high suggests that mothers are able to retrieve information like for example, number of recipes taught and their ingredients.

Skills

Being able to apply the information recalled goes a step further than thinking on how to apply it.

Both before and after adjusting for confounding, intervention II scores were significantly higher than intervention I. Results suggest that there is an added benefit of receiving the recipe demonstrations in an one-to-one setting. It seems that receiving the recipe demonstrations in an one-to-one setting improves mother's cooking skills and possibly her confidence in doing it. Chowdhury (1988)[96] also compared skills acquisition on the preparation of ORT between group and one-to-one approaches. The study found no significant differences between the two approaches when subjects were asked to prepare the ORT on their own. However, the class size in Chowdhury's study was only 5 while in the TSNI it was an average of 14 mothers in each class.

The AHRQ report[59] and Wrieden *et al* (2007) study[181] report on the importance of practical food skills intervention for teaching and improving nutrition/dietary change in developed countries. For example, in the former, findings show that Interventions including "interactions with food," such as cooking or taste testing, seemed particularly promising in increasing fruit and vegetable intake and reducing fat intake.

This research study has not found any other study reporting to have measured the impact of interventions to improve food preparation skills in developing countries with systematic direct observations. Cooking is a practical skill intervention, but the majority of studies have used questionnaires to assess it.

Block 3 output steps - Limitations and suggestions for future research:

The only limitation is that no sample size calculations were carried out. The sample size chosen was based solely on logistic and time limitations. It is possible that a bigger sample size would be needed to see a significant difference for comprehension-psychological meaning and recall between the two intervention groups.

6.3.4 Attitudes towards breastfeeding and complementary feeding practices

Whilst the previous output steps measure points in the persuasive process and measurements were collected while the TSNI project was on going, the attitude tool was administered when the project had already finished and hence measures an overall impact of the nutrition education program. This is not to say that attitudes are the most important measure. In fact, this study suggests that skills acquisition is the most important of all the measured output steps in the persuasion process. Because of the design of this study the attitude scale only compares treatment groups (i.e. intervention I, intervention II and control group). While with the previous output steps the setting from which they were measured could influence the results (i.e. the all one-to-one versus group setting discussion), attitudes were all measured in the same way. But this is not to say that attitude scores are not influenced by the delivery method that the mothers were exposed to during the nutrition education program.

The attitude tool in this study is a measure of mothers' agreement or non agreement with dietary practices (breastfeeding and complementary feeding, respectively) considered detrimental to children under five years. A control group was used as a proxy for attitude change, hence, this study compared both change in attitudes between the two intervention groups, but also positive attitudes between the two intervention groups.

For the section on breastfeeding practices, both intervention groups showed significantly higher positive attitudes than the control group, suggesting that the TSNI study improved mothers' attitudes towards breastfeeding practices. However, there was no significant difference between the two intervention groups. The main reason for this lack of difference could be due to the fact that while both intervention groups were exposed to group sessions on breastfeeding practices, the latter were not covered/taught in the individual classes, so in this case, there was not any added

benefit of receiving IC. However, we don't know what would be the result if breast feeding practices had also been taught in the individual classes.

Before adjusting for the education level of the mothers, intervention II mean scores were significantly higher than intervention I for the section on complementary feeding attitudes. After adjusting for the education level, intervention II mean scores were only marginally higher than intervention I. However, the difference between control and intervention group was only statistically significant for intervention II suggesting that mothers in intervention II gained more positive attitudes towards complementary feeding practices than intervention I. It would have been expected for attitudes to this section to be higher for intervention II since as argued by McGuire, the repetition of messages aids attitude change.[71]

There was a quite a difference in scores between the two sections of the attitude tool for the control group. While scores for the section on breastfeeding practices were very low, those for the section on complementary feeding practices were much higher. There could be two reason for this: first, that mothers seem to be more resilient in changing the more traditional/taboo practices, which was the case of the breast feeding section. This was also seen with the intervention groups where the practices with lower scores were those considered more traditional and more like taboos. This reasoning is also supported by Krueger and Gericke (2003)[185] who discuss that “cultural factors and taboos have a powerful influence on feeding practices and eating patterns. Young mothers often find it impossible to ignore their ill-informed elders or peer groups.” [185] This is also supported by Ashworth and Feachem (1985): “Mothers will not change their practices if entrenched cultural beliefs are contravened...”[22]

There is also another possible reason for why scores in section B for the control group were much higher. It is possible that control households became exposed to educational and promotional messages, regarding complementary feeding, that the TSNI project broadcasted via the provincial radio.

Another interesting finding is that attitudes in Malei locality was higher than Posto Campo/Luala and Catale. A suggestion for this difference is that probably the health promoter working in Malei was more active and energetic and possibly persuasive than all the others. However, because there are no baseline results for attitudes for each of these localities, it is not possible to know on which of these localities was the attitude change the biggest.

6.3.5 Degree of persuasiveness

As was seen previously, there were mixed results for the difference between the intervention groups for each of the output steps. Because we are breaking down the comparison by output step, it becomes difficult to reach a general conclusion about the effectiveness of the intervention groups. To make it possible for this generalisation, an overall factor, *persuasiveness degree*, was created.

Exposure, attention, comprehension, agreement, recall and skills acquisition are all points in the communication persuasion process that the receiver of the communication has to go through in order for the message to have an impact. All of them together make up the persuasiveness degree of a communication.

In this study, persuasiveness degree in each intervention group was measured as the sum of the values for each output step. To account for different units among the output steps, all values were standardised to proportions. Moreover, a weight was given to each output step according to their position and importance in the persuasive process. This allowed for more important output steps, like for example skills acquisition, to have a heavier weight on the equation than a weaker output step like for example exposure. A similar method has been used by Schellenberg *et al* (2003)[186] whereby they constructed an index of socio-economic status which combined household-level information on assets, income sources, and education. However, instead of allocating weights the way this research did, based on theory, they used the principal components analysis to define the weights. This PhD research study could not have used the principal component analysis simply because the several items (i.e. the output steps) in the index (i.e. the persuasiveness degree factor) were not measured in the same subjects.

Like, Schellenberg *et al* (2003)[186] index, which showed that the index's score explained 3% of the variation in height-for-age and was a better predictor of height-for-age than were any of the individual components of the score, the index developed in this study could also be used to explain behaviour change.

This research study has shown that the mean persuasiveness degree (or the index score) of intervention II was significantly higher than intervention I. Hence, even though for some output steps intervention II did not score significantly higher than intervention I, results suggest that intervention II's persuasiveness strength (measured as persuasiveness degree) was significantly higher than intervention I.

Results suggest that overall, the intervention II delivery method was more effective in persuading mothers up to the point before behaviour change. What remains to find out is if this higher effectiveness in intervention II would also lead to a significantly higher behaviour change.

6.3.6 Comparing TSNi results between the two intervention groups

Unfortunately, even though the most appropriate approach would have been to relate the current study output steps' results with TSNi results, this was not the case. The ideal design would have been to associate some of the output steps' results with results seen in the TSNi study population, especially for behaviour change, since McGuire's theory aims at influencing behaviour through the communication process. This was not done because behaviour change, as measured with participant observation during meal times, was only carried out for intervention II and hence, it would not be able to do a comparison between the two intervention groups. If the intervention was to be repeated again, a sub-group of subjects could have been randomly chosen from intervention I where meal time participant observation could also be carried out. As it is, there are plans to associate the results from the attitude tool with Food Frequency Questionnaires collected during the TSNi surveys which could give an idea of change in child feeding practices.

Published work on the TSNi study [106, 113] (see section 3.7) has also compared outcomes like for example mothers knowledge, reference child nutrient intake and nutritional status between the two intervention groups. The latter is not presented and discussed here since statistical analysis are still undergoing.

Mothers knowledge

To roughly quantify the change in nutritional knowledge over time, a simple index was created assigning 1 point for each correct answer. Overall there were 12 questions and therefore the maximum score was 12. Of these 12 questions, 4 referred to vitamin A knowledge, 5 to breastfeeding practices, and another 3 to complementary feeding practices. There was no significant difference in change (between baseline and end of study) between the intervention groups after adjusting for age, formal education levels and other characteristics of mothers and households. These results also support this research study's results for comprehension-psychological meaning where no difference between the groups was found. Interestingly, results for the TSNi knowledge index suggested that attending group sessions improved knowledge since for every

additional group session attended, all other factors being equal, the nutrition knowledge score increase by 0.30. [105] [106, 113]

Reference child nutrient intake

Results comparing reference child nutrient intakes between the two intervention groups at the end of the study showed no significant difference in the average intake levels of vitamin A. However, for the three key macro-nutrients (energy, protein and lipids) and many other micronutrients (iron, zinc, selenium, phosphorus, magnesium, folate, thiamine, riboflavin, niacin and vitamin B6) average intake levels were significantly higher for intervention II than for intervention I. The principal investigator in the TSNI study argued that “one plausible explanation for these differences is that the key messages promoting OFSP and other vitamin A rich food consumption in the household visits was simple to understand and relatively easy for households to adopt. Adopting the more difficult practices (for example, adding a small amount of fat into each main meal) may have required greater reinforcement and encouragement that more direct one-on-one contact provides.” [105] [106, 113]

As was discussed previously in section 3.6, chapter 3, the breadth of the TSNI intervention was extensive, putting it into the category of complex interventions. The nutrition education intervention was in itself extensive, covering a lot of messages, which makes it difficult to know which messages have contributed to which results. At this level, this array of messages might interact with each other, complicate and bias the outcomes of the intervention.

6.3.7 Which intervention group is better?

Care must be taken when interpreting this question. One intervention might be better than the other in terms of its effectiveness. As it was shown in this section, one intervention group was considered more effective than the other if it scored significantly (in statistical terms) better for the output steps. This was the case for intervention II.

In addition, one intervention might also be better than the other if it fulfils a certain criteria when the other does not. If for example, the criteria was attendance to 75% of total classes, then even if intervention II was more effective, if it did not reach that target attendance rate, it would not be considered as a better delivery method than intervention I. Like wise, intervention II might be more effective than intervention I, but intervention I might nonetheless reach all the criteria.

The importance of these criteria or standards has already been discussed in section 6.1.4.

An important criteria for this study, which was not measured, could have been certain changes in dietary practices, or even prevalence of certain measures indicating positive nutritional status of reference children (i.e. 75% with serum retinol above 20 $\mu\text{g dl}^{-1}$ or 75% with MUAC above cut-off point for age).

Nonetheless, instead of end of process outcomes like behaviour change and nutritional status, these criteria could instead be developed for each of the output steps.

While this study does not suggest per se criteria for the output steps and instead leaves this as a recommendation for future research, the following is an account of how the TSNi performed for each of the output steps, which can also be used as a starting point for the development of cut-off points for these standards or criteria.

Exposure

Attendance rates to the TSNi nutrition education program varied according to the approach. For the group setting approach (i.e. GC), 44% of mothers (regardless of intervention group) attended 10-12 classes. We cannot say if 44% is optimal as a cut-off point rate has not yet been defined, but as was seen was previously in section 6.3.6, for every additional group class attended, the nutrition knowledge score (developed by the TSNi study) increased by 0.30.[105] But what would be the optimal attendance rate? Perhaps having less than 50% of mothers attending 10-12 classes is sub-optimal.

For the one-to-one approach (i.e. IC), 87% of mothers attended 5-6 classes. Even without a cut-off point this proportion seems very high.

Block 2 output steps:

For the group class approach (i.e. GC), 56% of mothers (regardless of intervention group) paid attention in 8-10 observations. For the one-to-one approach (i.e. IC), 74.4% of paid attention in 8-10 observations. Even though there was definitely a higher proportion of mothers with high levels of attention in IC, we still do not know if this is the optimal level. Or perhaps, a proportion of 50% is enough and hence both approaches performed well.

Moreover, for the one-to-one approach 81.3% of the mothers answered correctly the recall questions. While, for the group class approach (regardless of intervention group) 65.0% of the mothers answered correctly the recall questions. Even without a cut-off point these proportion seem optimal.

Regarding the comprehension lexical meaning question, 61.2% of mothers in GC (regardless of intervention group) answered it correctly, and 71.4% of mothers in IC answered it correctly. Both proportions seem good, and if the optimal proportion was 60% both approaches would have performed well.

Block 3 output steps:

As would be expected, only around 30% of the mothers in intervention I and 35% in intervention II were able to answer correctly more than 70% of the comprehension-psychological meaning questions, which is low. So one can see that while for comprehension-lexical meaning, which consisted of relatively simpler questions, mothers were able to score highly. In turn, for comprehension-psychological meaning, which consisted of more complex questions, mothers found it more difficult to answer correctly the questions.

Around 63% of the mothers in intervention I and 75% in intervention II (were able to answer correctly 70-100% of the recall questions. This suggests that the TSNI study performance in increasing recall of certain messages, including those related to a food preparation task, was good.

Around 70% of mother in intervention II and 50% in intervention I were able to perform correctly 5-6 steps of the skills acquisition task. Even without a standard, this suggests that the TSNI study performance in increasing food preparation skills in intervention II mother's was good, but not so for intervention I mothers.

48.2% of intervention I mothers, 45% of intervention II mothers and 7.7% of control mothers had a score between 36-45 for the attitudes items regarding breastfeeding practices. Obviously that compared with the control group, the intervention groups did much better which suggests that the TSNI study was able to drastically improve attitudes towards breastfeeding practices. However, is a proportion less than 50% good?

44% of intervention I mothers, 56.4% of intervention II mothers and 37.1% of control mothers had a score between 36-45 for the attitudes items regarding complementary feeding practices. The difference between the intervention groups and control group was much smaller than for the breastfeeding section, moreover the proportion for intervention I was less than 50% which is sub-optimal.

It would have been good to be able to assess the TSNI nutrition education program in terms of the mothers performance for each output step. However, without standards it becomes difficult to do so and hence, one can only make arbitrary decisions on what is good performance.

6.3.8 Summary of findings for this section

The main finding in this section is that nutrition education messages delivered through intervention II are more effective in terms of their persuasive power as, overall, the persuasiveness degree of intervention II was stronger (i.e. statistically significantly higher) than intervention I.

The TSNI results on reference children dietary intake also support this finding as results for this outcome for intervention II was significantly higher than intervention I.

The main reason which explains higher significant results for intervention II is likely to be the added benefit of receiving individual classes. The processes in a one-to-one setting which are different from a group setting, seem to facilitate exposure to message, attention to message, comprehension of message, and later recalling of information. The women receiving one-to-one messages also seemed to be more easily persuaded into agreeing with the messages. Moreover, receiving demonstrations in an one-to-one setting seems to facilitate the acquisition of practical cooking/food preparation skills. One of the possible ways that makes it such a stronger approach is the possibility of tailoring the messages and the communication to the mothers' needs. The health promoter has to chance to understand the easiness and the barriers that the mothers have in adopting certain recommendations and then adapting it if necessary. Since it was the individual classes that seemed to have made a contribution to the strength of intervention II, and not necessarily the group classes (since group setting results for intervention II were not significantly higher than group setting results for intervention I whenever possible to compare the two), then it does seem that it was not necessary the use of the more costly combined approach employed by the TSNI

study. Instead the TSNI study should have only compared a GC only approach against a IC only approach. However, another study would be needed to compare these two approaches in the same kind of study setting, as IC was measured in this study along with the synergistic effect of being combined with a group setting approach, and the only way to break this synergistic effect is to have IC on its own.

This study suggest that the main processes that play a big role in improving persuasiveness in an one-to-one setting are: low disturbance levels, full attention given to the mother by the health promotes, which also means more control over the inattentiveness of the mother, more opportunity for feedback and discussion between health promoter and the mother. A report written by Graesser *et al* [187] suggests that the advantage in one-to-one school tutoring may be attributed to conversational dialogue patterns. When they explored this, they found out that students are more active, they tend to ask more questions, in tutoring contexts than in classroom settings. They also discuss that there is a balanced collaboration between the teacher/tutor and the student while they solve problems and answer questions, and also another advantage of tutoring is more preoccupation that the student is understanding and a dialogue between the two to ascertain this.

There are however main limitations found in the intervention II delivery method that need to be weighed against its benefits:

This study has found that even though there is an added benefit of receiving individual classes, attendance to individual classes also has a detrimental effect on the attendance to group classes. Moreover, the TSNI knowledge results show that attendance to group classes seem to have a significant impact on improvement of nutrition knowledge. It has also been shown that knowledge is an important (but not sufficient) factor predicting behaviour change.

Chowdhury study[96] calculated that the cost per mother receiving individual classes was twice than that for mothers receiving the intervention in a group setting. Neither this research study nor the TSNI study have carried out a cost-effectiveness evaluation of both intervention groups. However, the cost per mother receiving individual classes was certainly higher than the cost per mother receiving the group class.

One should also account for time constraint, as program coordinators will have to decide between delivering health messages to one women in a one-to-one setting or delivering to an average of 14 women in a group setting.

While it is possible to compare the two intervention groups in terms of their different performances for each output step, it is not possible to decide if the impact of each of the delivery methods on the performance for each output step is important as there are no standards to compare to.

6.4 Overall main study design limitations

Probably the one of the major limitations in this study which derives from a limitation in the study design of the TSNi study is that there was not a third arm, a individual class only intervention group. The comparison between the two intervention groups only allow us to make inferences on the added benefit of the IC approach, assuming that there is a synergistic effect of having both GC and IC together. But we don't know if a IC approach would be superior to a GC approach. We have seen in chapter 2 that studies that have compared both approaches found no differences between the two, however, outcomes in these studies were different from the outcomes in this research study. Moreover, if the IC approach proved to be more effective than the GC approach, there would be no point in combining both approaches as this would be the least cost-effective of the 3 interventions (i.e. IC only, or GC only, or GC plus IC).

Another major limitation is that the current study measured most of the outcomes in McGuire's communication process but did not measure the last one, behaviour change. Hence, whilst there is an overwhelming amount of process data the relevance of these findings can only be made clear by relating this to behaviour change ("the so what factor"), as is dictated by McGuire's theory. The ideal design for the current study should have included a measurement of behaviour change for both intervention groups which could then be associated with some of the previous output steps. As it is, it is only possible to compare the two intervention groups based on the communication process output steps but not based on the step that would measure the ultimate impact- behaviour change.

Another limitation in this study is that intervention I sample size for some of the output steps was smaller than calculated for. This could have made the study underpowered and rendered it insufficiently sensitive to detect statistically significant differences between the experimental conditions.

Another limitation to the study design of block 2 output steps' measurements in GC is the fact that the identification information of mothers was not collected. This made it

impossible to associate this data with other output steps. The reason why this was done so, is because at the time it seemed very impractical to collect this information as it would mean that mothers would be aware that they were being measured and this could change their behaviour, especially for the attention observations. A suggestion to overcome this problem in future research would be to wait for all the subjects to seat down, choose the subjects to be measured, certify that they are study subjects (non study subjects also attended the TSNi nutrition education classes) and then ask the names of the all class as if checking their attendance, and so permitting the identification of those to be measured.

While all the studies that have used McGuire's model have measure the output steps on the same subjects, this research study measured them in different subjects. The main limitation of this is that it made it very difficult to look at the associations among output steps. The main reason why this research study did not measure the output steps on the same subjects is because unlike the other studies, measurements were not taken on one occasion, but through out the nutrition education program. This allowed the measurement of different aspects of the program, like for example, behaviour and recall abilities in a group and individual class, other different learning outcomes in a practical exercise (i.e. the recipe task), and also, learning when the program had finished. If this research study had decided to measure the same subjects and follow them through out the project to be able to measure all output steps, the sample size would have needed to be much bigger to account for those mothers not attending classes or dropping out of the TSNi project.

Another limitation of this research study design is that not the same type and number of comprehension-lexical meaning questions were administered to the different intervention-approach groups. While the numbers were very similar when the comparison was between approaches, this similarity changed when results were instead broken down by intervention group.

6.5 CONCLUSIONS

This research study was the first one to have used McGuire's model and his multi-step sequence to measure communication/persuasion in a nutrition education program. It was also the first one to compare a group class only delivery method with a group class plus individual class delivery method in a developing country setting.

McGuire's communication/persuasion model emerged as a good conceptual framework to use in the (process) evaluation of nutrition education programs as it is a practical and user friendly model. But most importantly, it allowed for the breakdown of the communication persuasion process into steps that allowed for the comparison of the two delivery methods. The steps in turn were able to detect differences between the two delivery methods.

It was the first time that such tools were developed and used in this type of setting and since they were able to detect variability between the intervention groups, it suggests that they might also be useful for other similar studies in developing countries.

The group class plus individual class delivery method scored significantly better in the following output steps: exposure, attention (measured via recall questions), comprehension (lexical), skills acquisition and change in attitudes towards complementary feeding attitudes.

Group class plus individual classes intervention leads to a higher degree of persuasion than a group class only intervention. However, for mothers in developing countries which are time restricted, attending both individual classes and group classes is quite difficult.

Even though it was shown that adding home visits to group classes improves the persuasiveness of the message, further research is needed to measure cost-effectiveness of this delivery method against a group class only delivery method in a developing country setting. Nonetheless, should also measure cost-effectiveness of the group class only delivery method against a home visit only delivery method, since the latter on its own without being combined with a group class approach could possible lead to significantly higher persuasiveness against the group class only delivery method.

The adapted model that emerged from this research may be specific to the context in which it was developed. Further research would be necessary to determine the extent of generalisability. However, it has shown to be a useful conceptual framework as it was able to detect differences between the two intervention groups and pick up stages in the process of communication.

References

1. WHO, *Global strategy for infant and young child feeding*. 2003, WHO: Geneva.
2. Estatística, I.N.d., *Actualização das Projecções da População Total e por Área de Residência*. 2004, Instituto Nacional de Estatística. Maputo, Mocambique.
3. UNDP, *Education and Human Development: trajectory, lessons and challenges for the 21st century. National Human Development Report*, in *National Human Development Report*. 2000, UNDP: Mozambique.
4. UNDP, *Gender, Women and Human Development: An agenda for the future*, in *National Human Development Report*. 2001, UNDP: Mozambique.
5. UNICEF, *The State of the World's children*. 1998: New York.
6. Bender, A. and D. Bender, *Oxford Dictionary of food & nutrition*. 1995, United Kingdom.
7. WHO, *Indicators for assessing Vitamin A deficiency and their application in monitoring and evaluating intervention programmes*. 1996.
8. Sommer, A. and K. West, *Vitamin A deficiency: health, survival, and vision*. 1996, Oxford: University Press.
9. UNICEF, *Strategy for improved nutrition of children and women in Developing Countries*. 1990: New York.
10. WHO, *World Health Report - Health Systems: Improving performance*. 2000, WHO.
11. Mozambique, W.V., *Assistência aos serviços de saúde na provincia da Zambézia. Proposta de projecto*. 2001, World Vision Mozambique: Republica de Moçambique.
12. Estatística, I.N.d., *Moçambique, Inquérito Demográfico e de Saúde 2003. Demographic and Health Survey*. 2005, Instituto Nacional de Estatística: Maputo.
13. Saúde, M.d., *Inquérito Nacional sobre a deficiência de vitamina A, prevalência de anemia e malária em crianças dos 6-29 meses e respectivas mães*. 2003, Ministério da Saúde: Maputo.
14. WHO, *Measuring Change in Nutritional Status. Guidelines for assessing the nutritional impact of supplementary feeding programmes for vulnerable groups*. 1983: Geneva.
15. WHO, *Physical Status. The use and interpretation of anthropometry*., in *WHO technical report series, No. 854*. 1995, WHO.

16. Low J, T.D., *Towards Sustainable Nutrition Improvement in rural Mozambique: Assessing macro- and micro-nutrient malnutrition through new cultivars and new behaviours. Technical Report: First year of project activities: September 2002- August 2003.* 2003.
17. NATIONS, F.A.A.O.O.T.U., *Agriculture food and nutrition for Africa - A resource book for teachers of agriculture.* 1997, FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS: Rome.
18. Contento, I.R., et al., *The effectiveness of nutrition education and implications for nutrition education policy, programs, and research: a review of research.* . Journal of Nutrition Education, 1995. 27: p. 279-418.
19. Glanz, K., F. Lewis, and B. and Rimer, *The scope of health promotion and health education.*, in *Health behaviour and health education.*, K. Glanz, F. Lewis, and B. and Rimer, Editors. 1997: San Francisco: CA. p. 3-18.
20. Ewles, L. and I. Simnett, *What is Health promotion?*, in *Promoting health: a practical guide.*, L. Ewles and I. Simnett, Editors. 1996, Bailliere Tindall. p. 20-21.
21. Whitehead, D., *Health promotion and health education: advancing the concepts.* Journal of advanced nursing, 2004. 47(3): p. 311-320.
22. Ashworth, A. and R.G. Feachem, *Interventions for the control of diarrhoeal diseases among young children: weaning education.* Bulletin of the World Health Organization, 1985. 63(6): p. 1115-1127.
23. Cerqueira, M.T., *Nutrition education: A review of the nutrient-based approach* Food, Nutrition and Agriculture, 1991. 1(2/3).
24. R. Sheikholeslam, Z. Abdollahi, and F.N. Haghighi, *Managing nutritional programmes in developing countries.* Eastern Mediterranean Health Journal, 2004. 10(6).
25. Smith, J.L. and L.M. Lopez, *The application of theory and its relationship to program effectiveness in nutrition education research.* Journal of Nutrition Education, 1991. 23(2): p. 59-64.
26. Bhandari, N., et al., *An educational intervention to promote appropriate complementary feeding practices and physical growth in infants and young children in rural Haryana, India.* J Nutr, 2004. 134(9): p. 2342-8.
27. Prasad, B. and A.M. Costello, *Impact and sustainability of a "baby friendly" health education intervention at a district hospital in Bihar, India.* BMJ, 1995. 310(6980): p. 621-3.
28. Nakao, R.M., K.I. Kennedy, and G. Savina, *Breastfeeding education and infant health in the rural Philippines.* Ecol Food Nutr, 1992. 27: p. 115-26.

29. Morrow, A.L., et al., *Efficacy of home-based peer counselling to promote exclusive breastfeeding: a randomised controlled trial*. Lancet, 1999. **353**(9160): p. 1226-31.
30. Leite, A.J., et al., *Effectiveness of home-based peer counselling to promote breastfeeding in the northeast of Brazil: a randomized clinical trial*. Acta Paediatr, 2005. **94**(6): p. 741-6.
31. Labarere, J., et al., *Efficacy of breastfeeding support provided by trained clinicians during an early, routine, preventive visit: a prospective, randomized, open trial of 226 mother-infant pairs*. Pediatrics, 2005. **115**(2): p. e139-46.
32. Jakobsen, M.S., et al., *Promoting breastfeeding through health education at the time of immunizations: a randomized trial from Guinea Bissau*. Acta Paediatr, 1999. **88**(7): p. 741-7.
33. Haider, R., et al., *Breast-feeding counselling in a diarrhoeal disease hospital*. Bull World Health Organ, 1996. **74**(2): p. 173-9.
34. Haider, R., et al., *Effect of community-based peer counsellors on exclusive breastfeeding practices in Dhaka, Bangladesh: a randomised controlled trial [see comments]*. Lancet, 2000. **356**(9242): p. 1643-7.
35. Froozani, M.D., et al., *Effect of breastfeeding education on the feeding pattern and health of infants in their first 4 months in the Islamic Republic of Iran*. Bull World Health Organ, 1999. **77**(5): p. 381-5.
36. Davies-Adetugbo, A.A., et al., *Breast-feeding promotion in a diarrhoea programme in rural communities*. J Diarrhoeal Dis Res, 1997. **15**(3): p. 161-6.
37. Davies-Adetugbo, A.A., *Promotion of breast feeding in the community: impact of health education programme in rural communities in Nigeria*. J Diarrhoeal Dis Res, 1996. **14**(1): p. 5-11.
38. Coutinho, S.B., et al., *Comparison of the effect of two systems for the promotion of exclusive breastfeeding*. Lancet, 2005. **366**(9491): p. 1094-100.
39. Bhandari, N., et al., *Effect of community-based promotion of exclusive breastfeeding on diarrhoeal illness and growth: a cluster randomised controlled trial*. Lancet, 2003. **361**(9367): p. 1418-23.
40. Alvarado, M.R., et al., *Evaluation of a breastfeeding-support programme with health promoters participation*. Food and Nutrition Bulletin, 1996. **17**(1): p. 49-53.
41. Akram, D.S., M. Agboatwalla, and S. Shamshad, *Effect of intervention on promotion of exclusive breast feeding*. J Pak Med Assoc, 1997. **47**(2): p. 46-8.
42. Green, C.P., *Improving breastfeeding behaviors: evidence from two decades of intervention research*. 1999, LINKAGES.

43. Dewey, K.G., *Infant nutrition in developing countries: what works?* Lancet, 2005. 365(9474): p. 1832-4.
44. Walsh, C.M., A. Dannhauser, and G. Joubert, *The impact of a nutrition education programme on the anthropometric nutritional status of low-income children in South Africa*. Public Health Nutr, 2002. 5(1): p. 3-9.
45. Santos, I., et al., *Nutrition counseling increases weight gain among Brazilian children*. J Nutr, 2001. 131(11): p. 2866-73.
46. Salehi, M., et al., *Assessing the impact of nutrition education on growth indices of Iranian nomadic children: an application of a modified beliefs, attitudes, subjective-norms and enabling-factors model*. Br J Nutr, 2004. 91(5): p. 779-87.
47. Penny, M.E., et al., *Effectiveness of an educational intervention delivered through the health services to improve nutrition in young children: a cluster-randomised controlled trial*. Lancet, 2005. 365(9474): p. 1863-72.
48. Kilaru, A., et al., *Community-based nutrition education for improving infant growth in rural Karnataka*. Indian Pediatr, 2005. 42(5): p. 425-32.
49. Hotz, C. and R.S. Gibson, *Participatory nutrition education and adoption of new feeding practices are associated with improved adequacy of complementary diets among rural Malawian children: a pilot study*. Eur J Clin Nutr, 2005. 59(2): p. 226-37.
50. Guldán, G.S., et al., *Culturally appropriate nutrition education improves infant feeding and growth in rural Sichuan, China*. J Nutr, 2000. 130(5): p. 1204-11.
51. Brown, L.V., et al., *Evaluation of the impact of weaning food messages on infant feeding practices and child growth in rural Bangladesh*. Am J Clin Nutr, 1992. 56(6): p. 994-1003.
52. Contento, I.R., J.S. Randell, and C.E. Basch, *Review and analysis of evaluation measures used in nutrition education intervention research*. Journal of Nutrition Education and Behavior, 2002. 34(1): p. 2-25.
53. Nutbeam, D. and E. Harris, *Theory in a nutshell: a guide to health promotion theory*. 1999, Roseville: McGraw-Hill.
54. Contento, I.R., *The Effectiveness of Nutrition Education and Implications for Nutrition Education Policy, Programs, and Research: A Review of Research*. Journal of Nutrition Education, 1995. 27(6): p. 277-385.
55. Kreuter, M., V. Strecher, and B. Glassman, *One size does not fit all: the case for tailoring print materials*. Annals of Behavioral Medicine., 1999. 21: p. 276-283.
56. Yépez, G., *Communicating the HIV/AIDS risk to Hispanic populations: A review and integration*. Hispanic Journal of Behavioral Sciences., 1992. 14(4): p. 403-420.

57. Sabogal, R., M. G., and R. Otero-SAbogal, *Hispanic familism and acculturation: what changes and what doesn't?* Hispanic Journal of Behavioral Sciences., 1987. 9: p. 397-412.
58. West, J., et al., *Self-Efficacy predicts adherence to dietary sodium limitation in patients with heart failure.*, in *Unpublished manuscript, Stanford University*. 1999: Stanford, CA.
59. *Efficacy of Interventions To Modify Dietary Behavior Related to Cancer Risk.*, E.R.T. Assessment:, Editor. 2000, Agency for Healthcare Research and Quality: Rockville, MD.
60. Gillespie, A.H., *Applying communication theory in nutrition education research*. Journal of Nutrition Education, 1981. 13(1 (supplement)): p. S29-S33.
61. Stiff, J. and P. Mongeau, *Persuasive Communication*. Second ed. 2003, New York: The Guilford Press.
62. Taillard, M., *Persuasive communication: The case of marketing*. UCL Working Papers in Linguistics 2000. 12: p. 145-172.
63. Gillespie, A.H., *Communication theory as a basis for nutrition education*. Journal of the American Dietetic Association, 1987. 87(9 supplement): p. S44-S52.
64. Gillespie, A.H. and P. Yarbrough, *A conceptual model for communicating nutrition*. Journal of Nutrition Education, 1984. 16(4): p. 168-172.
65. Yarbrough, P., *Communication Theory and Nutrition Education Research*. Journal of nutrition education, 1981. 13(1 supplement): p. S16-S26.
66. Bandura, A., *Social learning theory*. . 1977, New Jersey Prentice-Hall, Inc.
67. Santrock, J., *Educational psychology*. 2nd ed. 2004: McGraw Hill.
68. Flay, B., D. DiTecco, and R. Schegel, *Mass media in health promotion: an analysis using an extended information-processing model*. Health Education Quarterly, 1980. 7(2): p. 127-147.
69. McGuire, W., *Public Communication as a strategy for inducing health-promoting behavioural change*. . Preventive Medicine, 1984. 13: p. 299-319.
70. McGuire, W., *Personality and attitude change: an information processing theory.*, in *Psychological foundation of attitudes.*, A. Greenwald, T. Brock, and T. Ostrom, Editors. 1968, Academic Press: New York. p. 171-196.
71. McGuire, J.W., *Attitudes and Attitude change.*, in *Handbook of Social Psychology*, G. Lindzey and E. Aronson, Editors. 1985, Lawrence Erlbaum Associates, Inc.: New York.
72. McGuire, W.J., *Chapter 4: The Communication/Persuasion Matrix.*, in *Evaluating Advertising: A Bibliography of the Communications Process B*.

- Lipstein and W.J. McGuire, Editors. 1978, Advertising Research Foundation: New York.
73. McGuire, W., *Theoretical foundations of campaigns*, in *Public Communication Campaigns*, R. Rice and C. Atkin, Editors. 1989, Sage: Thousand Oaks, CA.
 74. McGuire, W., *Input and Output variables currently promising for constructive persuasive communications*, in *Public communication campaigns*, R. Rice and C. Atkin, Editors. 2001, SAGE publications: California. p. 22-48.
 75. Kreuter, M. and S. McClure, *The role of culture in health communication*. Annual Review of public health, 2004. **25**: p. 439-55.
 76. Houts, P., et al., *The role of pictures in improving health communication: A review of research on attention, comprehension, recall and adherence*. Patient Education and Counseling, 2006. **61**: p. 173-190.
 77. Eagly, A.H. and S. Chaiken, *The psychology of attitudes*. 1993, New York: Harcourt, Brace Jovanovich.
 78. Rimer, B.K. and B. Glassman, *How do persuasive health messages work? A health education field study*. Health Education Quarterly, 1984. **11**(3): p. 313-321.
 79. Thomas, P.R., *Improving America's diet and health: From recommendations to Action*, in *A report of the Committee of Dietary Guidelines Implementation*. 1991, Food and Nutrition Board. Institute of Medicine.
 80. Greenhalgh, T., *Assessing the methodological quality of published papers*. BMJ, 1997. **315**(7103): p. 305-308.
 81. Houts, P., et al., *Using pictographs to enhance recall of spoken medical instructions*. Patient Education and Counseling, 2001. **43**(3): p. 231-242.
 82. Elder, J., *Monitoring and Evaluation plan for BCC activities PCI-Zambia*, Editor. 2004, Unpublished.
 83. Hausser, D., et al., *Evaluation des campagnes nationales de prevention contre le SIDA: Modele d'analyse*. Sozial und Praventivmedizin, 1987. **32**((4-5)): p. 207-9.
 84. Navaie, M., D. Glik, and K. Saluja, *Communication effectiveness of postnatal nutrition education in a WIC program*. Journal of Nutrition Education, 1994. **26**(5): p. 211-7.
 85. Russon, C. and L. Koehly, *Construction of a scale to measure the persuasive impact of qualitative and quantitative evaluation reports*. Evaluation and Program Planning, 1995. **18**(2): p. 165-177.
 86. Comino, E., et al., *The Australian National Asthma Campaign: effects of public education activities based on mass media*. American Journal of Preventive Medicine, 1997. **13**(4): p. 251-256.

87. Bull, F., *Understanding the effects of printed health education materials: which features lead to which outcomes?* Journal of health communication 2001. 6: p. 265-279.
88. Kreuter, M., et al., *Responses to behaviourally vs culturally tailored cancer communication among African American women.* American Journal of Health Behaviour, 2004. 28(3): p. 195-207.
89. Zhao, Y. and J. Ramaprasad. *The effect of Fear Appeals in AIDs prevention Ads on Attention, Interest, Liking and Intent to Adopt recommended behaviour.* in Association for Education in Journalism and Mass Communication 2005. San Antonio, Texas.
90. Ruiter, R., L. Kessels, and B. Jansma, *Increased attention for computer-tailored health communication: An event-related potential study.* Health Psychology, 2006. 25(3): p. 300-306.
91. Mensing, C. and S. Norris, *Group education in Diabetes: effectiveness and Implementation.* Diabetes Spectrum, 2003. 16: p. 96-103.
92. Rickheim, P.L., et al., *Assessment of group versus individual diabetes education: a randomized study.* Diabetes Care, 2002. 25(2): p. 269-74.
93. Hakala, P., R.L. Karvetti, and T. Ronnema, *Group vs. individual weight reduction programmes in the treatment of severe obesity--a five year follow-up study.* Int J Obes Relat Metab Disord, 1993. 17(2): p. 97-102.
94. Siero, F.W., et al., *Impact of group nutrition education and surplus value of Prochaska-based stage-matched information on health-related cognitions and on Mediterranean nutrition behavior.* Health Educ Res, 2000. 15(5): p. 635-47.
95. Rodriguez-Garcia, R., K.J. Aumack, and A. Ramos, *A community-based approach to the promotion of breastfeeding in Mexico.* Journal of Obstetric, Gynecologic, and Neonatal Nursing, 1990. 19(5): p. 431-8.
96. Chowdhury, A.M., F. Karim, and J. Ahmed, *Teaching ORT to women: individually or in groups?* J Trop Med Hyg, 1988. 91(6): p. 283-7.
97. Cormier, A., *Group versus individual dietary instruction in the treatment of obesity.* Can J Public Health, 1972. 63(4): p. 327-32.
98. Adams, S., K.E. Grady, and C. Mukaida, *Weight loss: A comparison of group and individual interventions.* Journal of the American Dietetic Association, 1986. 86(4): p. 485--489.
99. D'Erasmus-Melkus., G.A., J. Wylie-Rosett, and J.A. Hagan, *Metabolic impact of education in NIDDM.* Diabetes Care, 1992. 15(7): p. 864--869.
100. Jeffery, R.W., et al., *Weight and sodium reduction for the prevention of hypertension: a comparison of group treatment and individual counseling.* Am J Public Health, 1983. 73(6): p. 691-3.

101. De Mello, E.D., V.C. Luft, and F. Meyer, *[Individual outpatient care versus group education programs. Which leads to greater change in dietary and physical activity habits for obese children?]*. J Pediatr (Rio J), 2004. 80(6): p. 468-74.
102. Moe, E.L., et al., *Promoting Healthy Lifestyles: Alternative Models' Effects (PHLAME)*. Health Educ Res, 2002. 17(5): p. 586-96.
103. Llorca, D., et al., *Educacion grupal frente a individual en pacientes diábeticos tipo 2. Atención primaria: Publicación oficial de la Sociedad Española de Familia y Comunitaria*, 2003. 32(1): p. 36-41.
104. Gucciardi, E., et al., *Assessment of two culturally competent diabetes education methods: individual versus individual plus group education in Canadian Portuguese adults with type 2 diabetes*. Ethn Health, 2007. 12(2): p. 163-87.
105. Low J, *Towards Sustainable Nutrition Improvement in rural Mozambique: Assessing macro- and micro-nutrient malnutrition through new cultivars and new behaviours. Key findings*. 2005.
106. Low JW, et al., *A food-based approach introducing orange-fleshed sweet potatoes increased vitamin A intake and serum retinol concentrations in young children in Mozambique*. Journal of Nutrition, 2007. 137(5): p. 1320-1327.
107. Ajzen, I., *Persuasive communication theory in social psychology: A historical perspective.*, in *Human behavior: Theory and applications in recreation, tourism, and natural resource management.*, M.J. Manfredi, Editor. 1992, Sagamore Publishing. : Champaign, IL.: p. 1-27.
108. Dickin, K., et al., *Designing by dialogue: a program planners' guide to consultative research for improving young child feeding*. 1997, Academy for Educational Development;: Washington, D.C.:
109. Ley, P., *Communicating with patients: improving communication, satisfaction and compliance*. Psychology and Health series 4, ed. D. Marcer. 1988: Chapman & Hall.
110. Abusabha, R., J. Peacock, and C. Achterberg, *How to make nutrition education more meaningful through facilitated group discussions*. J Am Diet Assoc, 1999. 99(1): p. 72-6.
111. *A FRAMEWORK FOR DEVELOPMENT AND EVALUATION OF RCTS FOR COMPLEX INTERVENTIONS TO IMPROVE HEALTH*. 2000, Medical Research council (MRC) UK.
112. Campbell M, Fitzpatrick R, and H. A., *Framework for design and evaluation of complex interventions to improve health*. BMJ, 2000. 321: p. 694-696.
113. Low JW, et al., *Ensuring the supply of and creating demand for a biofortified crop with a visible trait: lessons learned from the introduction of orange-fleshed sweet potato in drought-prone areas of Mozambique*. Food and Nutrition Bulletin, 2007. 28 ((2 Suppl)): p. S258-70.

114. Cohen, L., L. Manion, and K. Morrison, *Research methods in education*. 5th ed. 2003: Routledge.
115. Miller, D.C. and N. Salkind, *Handbook of Research Design and Social Measurement* 6th ed, ed. T. Oaks. 2002, California: Sage.
116. Streiner, D.L. and G.R. Norman, *Health measurement scales: a practical guide to their development and use*. 3rd ed. 2003, New York: Oxford University Press.
117. Reynolds, C.R., R.B. Livingston, and V. Willson, *Measurement and assessment in education*. 2006: Pearson Education, Inc.
118. Kirkwood, B.R., *Essentials of Medical Statistics*. 2002, Oxford: Blackwell Science Ltd.
119. Branson, B., et al., *Group counselling to prevent sexually transmitted disease and HIV: A randomized controlled trial*. *Sexually transmitted diseases*, 1998. **25**(10): p. 553-560.
120. James, W., *Principles of psychology*. 1890, New York: Henry Holt
121. Zeaman, D. and B. House, *The role of attention in retarded discrimination learning*. , in *Handbook of mental deficiency*., N. Ellis, Editor. 1963, McGraw-Hill: New York. p. 159-233.
122. Boersma, F. and W. Muir, *Eye movements and information processing in mentally retarded children*. 1975, Rotterdam: Rotterdam University Press.
123. Piontkowski, D. and R. Calfee, *Attention in the classroom* in *Attention and Cognitive Development*, G. Hale and M. Lewis, Editors. 1979, Plenum Press: New York and London.
124. Lee, S., *Encyclopedia of school psychology*. 2005, California: SAGE publications.
125. DeGangi, G. and S. Porges, *Neuroscience Foundations of Human Performance*. 1990, Rockville, MD: American Occupational Therapy Association Inc.
126. www.wikipedia.org. [cited.
127. Salvucci, D. *A model of eye movements and visual attention*. in *Proceedings of the International Conference on Cognitive Modeling*. . 2000. Veenendaal, The Netherlands: Universal Press.
128. Smith, D.T., C. Rorden, and S.R. Jackson, *Exogenous Orienting of Attention Depends upon the Ability to Execute Eye Movements*. *Current Biology*, 2004. **14**: p. 792-795.
129. Yarbus, A.L., *Eye movements during perception of complex objects*., in *Eye Movements and Vision*., L.A. Riggs, Editor. 1967, Plenum Press: New York. p. 171-196.

130. Corbetta, M. and e. al., *A Common Network of Functional Areas for Attention and Eye Movements*. Neuron, 1998. **21**: p. 761-773.
131. Fischer, M., et al., *Recall and eye tracking study of adolescents viewing tobacco advertisements*. JAMA, 1989. **261**(1): p. 84-89.
132. Kaskutas, L. and T. Greenfield, *The role of health consciousness in predicting attention to health warning messages*. . American Journal of Health Promotion, 1997. **11**(3): p. 186-193.
133. Delp, C. and J. Jones, *Communicating information to patients: the use of cartoon illustrations to improve comprehension of instructions*. Academic Emergency Medicine, 1996. **3**: p. 264-270.
134. Cole, R. and S. Holland, *Recall of health education display materials*. Health Education Journal, 1980. **39**(3): p. 74-9.
135. Peterson, P.E., et al., *How pronutrition television programming affects children's dietary habits*. Developmental Psychology, 1984. **20**(1): p. 55-63.
136. Palfrey, J.S., et al., *An analysis of observed attention and activity patterns in preschool children*. Journal of Pediatrics, 1981. **98**(6): p. 1006-1011.
137. Kingston , A., *Cognitive Ethology: A New Approach for Studying Human Cognition*. p. Editorial Manager(tm) for British Journal of Psychology Manuscript Draft
138. www.wikipedia.com. [cited.
139. Bransford, J., *Human cognition: learning, understanding and remembering*. 1979, Belmont, California: Wadsworth Publishing company.
140. White, R. and R. Gunstone, *Probing Understanding*. 1992, London: The Falmer Press.
141. Yingxu, W. and D. Gafurov. *The cognitive process of comprehension*. in *Cognitive Informatics. Proceedings. The Second IEEE International Conference on*. 2003.
142. Long, D., *Second language listening comprehension: A Schema-Theoretic Perspective*. Modern Language Journal, 1989. **73**(1): p. 32-40.
143. Kools, M., et al., *Increasing readers' comprehension of health education brochures: a qualitative study into how professional writers make texts coherent*. Health education and Behaviour, 2004. **31**(6): p. 720-40.
144. Cheng, H., *Understanding in Education*, in *Institute of Education*. 1995, University of London: London.
145. Reeder, E. *Measuring what counts: memorization versus understanding*. 2002 [cited 2002.

146. Deyes, T., *Towards an authentic 'discourse cloze'*. *Applied Linguistics*, 1984. 5(2): p. 128-137.
147. Perez-Cuevas, R., et al., *Immunization promotion activities: are they effective in encouraging mothers to immunize their children?* *Social Science & Medicine* 1999. 49: p. 921-932.
148. Curtis, V., S. Cairncross, and R. and Yonly, *Review: Domestic hygiene and diarrhoea – pinpointing the problem*. *Tropical Medicine and International Health*, 2000. 5(1): p. 22-32.
149. VanLehn, K., *Cognitive Skills acquisition* *Annual Review of Psychology*, 1996. 47: p. 513–39.
150. Nagarajan, L., et al., *A comparative study of different methods of training of rural subjects for reconstitution of oral rehydration solutions*. *Indian Pediatr*, 1989. 26(4): p. 323-9.
151. Taylor, T., et al., *Knowledge, skills, and behavior improvements on peer educators and low-income hispanic participants after a stage of change-based bilingual nutrition education program*. *Journal of community health*, 2000. 25(3): p. 241-261.
152. Anderson, A., et al., *A questionnaire assessment of nutrition knowledge - validity and reliability issues*. *Public Health Nutrition*, 2002. 5(3): p. 497 - 503.
153. Giere, S., E. Rudrod, and S. and McKay, *A functional approach to meal planning, preparation, and shopping skills* *Australia and New Zealand Journal of Developmental Disabilities*, 1989. 15(2): p. 81-97.
154. Peterson, L., *Verbal Learning and Memory*. *Annual Review of Psychology*, 1977. 28: p. 393-415.
155. Atkinson, R. and R. Shiffrin, *Human memory: A proposed system and its control processes*. , in *The psychology of learning and motivation: Advances in research and theory*., K. Spence and J. Spence, Editors. 1968, Academic press: New York.
156. Raaijmakers, J. and R. Shiffrin, *Models for recall and recognition* *Annual Review of Psychology*, 1992. 45: p. 205-34.
157. Richardson-Klavhen, A. and R. Bjork, *Measures of memory*. *Annual Review of Psychology*, 1988. 39: p. 475-543.
158. Patel, V., T. Eisemon, and J. Arocha, *Comprehending instructions for using pharmaceutical products in rural Kenya*. . *Instructional Science*, 2004. 19(1).
159. Fishbein, M. and I. Ajzen, *Belief, attitude, intention, and behaviour: An introduction to theory and research*. . 1975.
160. Petty, R.E., D.T. Wegener, and L. Fabrigar, *Attitudes and attitude change*. *Annual Review of Psychology*, 1997. 48: p. 609-47.

161. Orstead, C., et al., *Efficacy of prenatal nutrition counseling: weight gain, infant birth weight, and cost-effectiveness*. J Am Diet Assoc, 1985. **85**(1): p. 40-5.
162. Cooper, J. and R. Croyle, *Attitudes and Attitude change*. Annual Review of Psychology, 1984. **35**: p. 395-426.
163. Schuman, H. and M. Johnson, *Attitudes and Behaviour*. Annual Review of Sociology, 1976. **2**: p. 161-207.
164. Hayes, N., *Principles of social psychology*. . 1993.
165. Hogg, M.A. and G.M. Vaughan, *Social psychology* 2002.
166. Hewstone, M. and W. Stroebe, *Introduction to social psychology*. 3rd ed. 2001: Blackwell.
167. Tesser, A. and D.R. Shaffer, *Attitudes and attitude change*. Annual Review of Psychology, 1990. **41**: p. 479-523.
168. Wallace, D.S., et al., *Which Behaviors Do Attitudes Predict? Meta-Analyzing the Effects of Social Pressure and Perceived Difficulty*. Review of General Psychology, 2005. **9**(3): p. 214-227.
169. Sax, G., *Principles of educational and psychological measurement and evaluation*. 3rd ed. 1989, Belmont, California: Wadsworth Publishing Company.
170. Kirkwood, B.R. and A.C. Sterne, *Essential Medical Statistics*. 2nd ed. 2003, Oxford: Blackwell Science.
171. Killip, S., Z. Mahfoud, and K. Pearce, *What is an intraclass correlation coefficient? Crucial concepts for primary care researchers*. Annals of family medicine, 2004. **2**(3).
172. Kerry, S.M. and J.M. Bland, *Analysis of a trial randomised in clusters*. BMJ, 1998. **316**(7124): p. 54.
173. Simpson, J.M., N. Klar, and A. Donner, *Accounting for Cluster Randomisations: A review of primary prevention trials, 1990 through 1993*. American Journal of Public Health, 1995. **85**(10): p. 1378-1383.
174. Field, A., *Discovering Statistics: using SPSS for Windows*. 2000, London: SAGE.
175. George, D. and P. Mallery, *SPSS for Windows step by step: a simple guide and reference*. 4th ed. 2003: Pearson Education, Inc.
176. Cohen, J., *Statistical power analysis for the behavioral sciences*. 2nd ed. 1988: Hillsdale, NJ: Lawrence Erlbaum Associates.
177. Santos, J.R.A., *Cronbach's Alpha: A tool for assessing the reliability of scales*. Journal of Extension, 1999. **37**(2).

178. U.S Census Bureau, M.I., Serpro S.A, *CSPro 3.0 Designer*. 20 Jan 2006.
179. Blatchford, P., *The class size debate. Is small better?* 2003: Open University Press.
180. Jacoby, E.R., et al., *Effectiveness of two methods of advising mothers on infant feeding and dietetic management of diarrhoea at an outpatient clinic in Peru*. Journal of Diarrhoeal Diseases Research, 1994. **12**(1): p. 59-64.
181. Wrieden, W., et al., *The impact of a community-based food skills intervention on cooking confidence, food preparation methods and dietary choices - an exploratory trial*. Public Health Nutrition, 2007. **10**(2): p. 203-211.
182. Liquori, T., et al., *The cookshop program: outcome evaluation of a nutrition education program linking lunchroom food experiences with classroom cooking experiences*. Journal of nutrition education, 1998. **30**(302-13).
183. Andrien, M., *Evaluation du programme d'education nutritionnelle du Centre Hospitalier Regional de Bouake*. . 1983, Laboratoire de Pedagogie Experimentale de l'Universite de Liege: Liege.
184. McDonald, D.D., M. Wiczorek, and C. Walker, *Factors affecting learning during health education sessions*. Clinical Nursing Research, 2004. **13**(2): p. 156-67.
185. Krueger, R. and G. Gericke, *A qualitative exploration of rural feeding and weaning practices, knowledge and attitudes on nutrition*. Public Health Nutrition 2003. **6**(2): p. 217.
186. Schellenberg, J.A., et al., *Inequities among the very poor: health care for children in rural southern Tanzania*. The Lancet, 2003. **361**.
187. Graesser, A.C., N.K. Person, and J.P. Magliano, *Collaborative Dialogue Patterns in Naturalistic One-to-one Tutoring*, Department of Psychology. The University of Memphis, Memphis.

APPENDICES

Appendix 4.1 Tool used to measure attention observations, attention recall and comprehension-lexical meaning in group classes – an example

Attention and understanding- group classes

Date: 04/06/04; **Loc:** 01 **Village:** 01 **Intervention:** 1

Farmer group: Namitudine

Topic of class: Hygiene Practices

No. total adults present: 14

No. total adult men present: 6

Duration of class : 40 min

Weight monitoring before? Yes

	Mothers					
<u>Observations*:</u>	1 st	2 nd	3 rd	4 th	5 th	6 th
1 st	0 (1)	1	1	1	1	1
2 nd	1	1	1	1	1	1
3 rd	1	1	1	1	1	1
4 th	1	1	1	1	1	1
5 th	1	1	1	1	1	1
6 th	0 (1)	1	1	1	1	1
7 th	0	0 (1)	1	1	0	1
8 th	1	1	1	0	1	0
9 th	1	0	1	0	1	0
10 th	1	1	0	1	1	1
Attention recall answers correct?	1 / 1	0 / 1	1 / 1	0 / 0	0 / 0	1 / 1

*Paid attention at the observation? Yes= 1 No= 0;

Answer to recall question correctly 1 = yes; 0 = No; 2= Does not remember

Child disturbing mother during direct observations? Yes - (1)

1st recall question: *Since I have started the session, did you hear me mention lepra?*

2nd recall question: *Since I have started the session, did you hear me mention undernutrition?*

Understanding question: where do the pathogens that cause diarrhoea come from?

	1 st mother	2 nd mother	3 rd mother	4 th mother	5 th mother	6 th mother
Answered correctly?	1	1	1	1	1	0
Yes -1 No-0						
If no, what answered?						Said water

Extra observations:

Total number of children crying during class: 7

Total number of women standing up during class: 6

Cont. Appendix 4.1

Health promoter performance

1. Speed of speech: very quick quick ☒ normal slow very slow

2. Loudness of voice very loud loud ☒ Normal low very low

3. Proportion of the audience covered by the health promoter:

All 1/2 1/3 1/4 ☒ 2/3

4. Number of questions asked: 3

5. Where questions asked by different people? ☒ Yes No

6. Did any of the women ask questions? ☒ Yes No

7. Attitude towards the audience:

Very friendly ☒ friendly indifferent not so friendly not friendly

9. Did the health promoter notice do something when she noticed that women were not paying attention?

☒ Yes No

10. What did she do? She told the women to pay attention

Appendix 4.2 Tool used to measure attention observations, attention recall and comprehension-lexical meaning at individual classes– an example

Attention and understanding measurements – home visits

Date: 04/06/04; Loc: 01 village: 31 Household: 21

Visit no.: 6

Total adults present: 1

Was the child with the mother? Yes

Duration: 30 min

Visit with observation? Yes No

Total number of men present: the husband
No

Observations*:		Observations:
	Key recommend:	
1 st	1 st	0 (1)
2 nd	2 nd	1
3 rd	3 rd	1
4 th	4 th	1
5 th	5 th	1
6 th	6 th	0 (1)
7 th	8 th	0
8 th	10 th	1
9 th	11 th	1
10 th	12 th	0 (1)
Attention recall answers correct?		1 / 1

*Paid attention at the observation? Yes= 1 No= 0;
Answer to recall question correctly 1 = yes; 0 = No; 2= Does not remember
Child disturbing mother during direct observations? Yes - (1)

1st recall question: Since I have started the session, did you hear me mention mushrooms?

2nd recall question: Since I have started the session, did you hear me mention dark green leaves?

Understanding questions:

- 1. Where do the pathogens that cause diarrhoea come from?
- 2. Overall, how many main meals should this child have? *Point to the reference child*
- 3. Give examples of 3 vitamin-A rich foods

1. Did she answer the understanding question correctly? Yes= 1 No= 0

Observations:

There was a lot of noise coming from neighbours but even so the mother was making eye contact

Health promoter performance

1. Speed of speech: very quick quick **normal** slow very slow
2. Loudness of voice very loud loud **Normal** low very low
3. Attitude towards the mother:
Very friendly **friendly** indifferent not so friendly not friendly
4. Did she ever noticed that the mother was not paying attention?
- Yes **No**

If Yes, what did the health promoter do?

5. Were they seating near to each other?
- Yes** No
6. Were they facing each other?
- Yes **No** they were seating side by side

Appendix 4.3 Tool used to measure comprehension-psychological meaning, skills acquisition and recall – an example

Date: 10/11/04 Locality: 02 Village: 28 HH: 07 Intervention: 1
Date of demonstration: 19/07/04 Starting time of task: 13:00 End time: 14:37

Ingredients used in today’s recipe:

- | | |
|--------------------------------|------------|
| 1. Orange fleshed sweet potato | 4. Coconut |
| 2. Fish | 5. _____ |
| 3. Dark green leaves | 6. _____ |

A. Cooking method used :

1. What was the cooking method used in today’s recipe?

1. Steaming 2. Normal cooking (i.e. boiling) 3. Other: _____

2. If it was the steaming method, did she perform well the following steps:
If not, how did she get it wrong? Describe.

- Add just a bit of liquid: she added too much water. And she boiled the water before.
- Cover with the lid when the pan is on the fire: she covered.
- If while cooking it is necessary to add more liquid then add only a bit: she did not have to add any more liquid as she had already added too much
- Do not uncover the pan for a long time! Only to give enough time to check if the OFSP is ready and if it needs more stirring : 1st uncovered to check and quickly covered; 2nd uncovered, stirred, and quickly covered; 3rd uncovered to check and quickly covered; 4th uncovered to check and quickly covered; 5th uncovered to check and took it off the fire.
- The OFSP should be thin sliced: she did
- Do not overcook the OFSP: OFSP cooked for 18 min and was very soft.
- Wash the child’s hand with something else apart from water: washed the child hands only with water
- Use dried coconut or dry it afterwards: she kept the grated coconut to dry it

Appendix 4.3 cont.

3. *If it was the boiling method, did she perform well the following steps:
If not, what steps and how did she get it wrong? Describe them.*

- Add enough water to cover the OFSP only to the middle:
- Should boil the water before adding the OFSP:
- Should chose potatoes of the same size, or cut the bigger ones, or even put the bigger ones below the smaller ones:
- Should not peel the OFSP:
- Do not overcook the OFSP:
- Should keep the cooking water to give the child to drink:
- Keep the grated coconut for drying:
- Wash the child's hand with something else apart from water:

B. Questions regarding the ingredients:

1st ingredient: BDPA

To which food group does it belong to?

She pointed to the correct food group

b) Why is this food group good for the child? Helps him grow

2nd ingredient: Dark green leaves

To which food group does it belong to?

She pointed to the fat rich food group

b) Why is this food group good for the child? Helps him grow

c) *If the food belongs to the protein or fat- rich groups then, what other ingredient from the same group could you use if you could not get this one?*

Appendix 4.3 cont.

3rd ingredient: fish To which food group does it belong to?

She pointed to the correct food group

b) Why is this food group good for the child? Helps him grow

c) *If the food belongs to the protein or fat- rich groups then, what other ingredient from the same group could you use if you could not get this one?*

Could use cashew nuts

4th ingredient: coconut To which food group does it belong to?

Pointed to the protein rich food group

b) Why is this food group good for the child? Helps him grow

c) *If the food belongs to the protein or fat- rich groups then, what other ingredient from the same group could you use if you could not get this one?*

Could use cashew nuts

1. How many recipes were taught in demonstration? 1 2 **3** 4 5

3. What were the ingredients used in each one?

1st recipe : OFSP – toasted peanut -- _____ -- _____ -- _____

2nd recipe: OFSP – fish – dark green leaves – coconut -- _____

3rd recipe: OFSP – coconut milk -- _____ -- _____ -- _____

4. Which one is the best for the child?

The one with fish

a) Why?

Because it has OFSP, coconut, fish and dark green leaves

5. *If the mother did not chose the recipe which she taught to be the best for the child, ask her why?*

Appendix 4.4 Attitude tool – an example

Loc: 06 Village: 62 HH: 03 Name: Teresinha Mendes Date: 07 / 12 /04

Duration: 11:59 – 12:15 Intervention: 1 2 Control

Has been questioned?: Yes No Type: 1 2 3

Likert Scale: (1) Strongly Disagree ☹️; (2) Disagree 😐; (3) Not sure 😐; (4) Agree 😊; (5) Strongly Agree 😊

Section A. Breastfeeding:

1. During those first 3 days when Maria was born, her grandmother would squeeze her mother’s breast to release that first yellow liquid that comes out. This liquid would then be discarded.

1

Because that 1st milk causes illness in the baby

2. Maria’s mother then prepared a tea made of leaves and roots given by the *curandeiro*. Maria was then given this tea.

1

Because that tea causes illness in the baby’s body

3. When Maria was 3 months old she got a really bad diarrhea. At that time her mother noticed that if she breastfed less often the baby would also have less diarrhea, therefore she decided to decrease the frequency of breastfeeding.

2

Because a baby with diarrhea needs to breastfeed very much

4. It is now January and it is raining a lot. Maria’s mother is always working the field. She goes as soon as the sun rises and only returns when the sun goes down. She wants to take advantage of this good rain to start sowing. But she is becoming tired of taking along Maria to the field. Because Maria is now 6 months old she has decided to stop breastfeeding her so that she could leave her at home with the grandmother.

1

Because the baby is still very small

5. Amélia is a month pregnant and also has a 6 month old baby that she still breastfeeds.

Because a pregnant mother cannot breastfeed her child

1

6. Now Amélia is 4 months pregnant. Her baby is now 9 months old. Therefore Amélia decided to stop breastfeeding her.

5

Because it can cause illness in the baby

7. Isabel got severely ill with malaria when her baby boy was 3 months old. She wanted to stop breastfeeding him but her sister, Teresa, told her not to do that. Therefore Isabel kept on breastfeeding even though she had malaria.

5

An ill mother can still breastfeed her child

8. Six months went by and Teresa went on a trip to Namacurra. At this same time Isabel got sick again with malaria. Her baby boy was already 9 months old. Without her sister's advice Isabel decided that it would be better to stop breastfeeding her baby boy.

5

Because the child is big

9. Teresa was away in Namacurra for 3 days. She left her own 9 month old child with her sister Isabel. When she returned from the trip she decided that she should not breastfeed him anymore.

5

Because the milk is already rotten

Section B - Complementary feeding:

10. It is enough for a crawling baby to eat twice a day.

1

Because it is not enough for a child that is crawling

11. The mother should not add an egg to the gruel of 6 month old baby.

1

A 6-month baby can eat egg

12. If a small child refuses to eat then the mother should not insist and just wait until he/she gets hungry.

2

Has to insist so that the child eats

13. The best gruel for the child is that one made of cassava flour and sugar as it fills the stomach.

(If subject's attitude is 1 or 2 write down the answer even for questionnaire type 2 and 3)

Because these porridges have no vitamin

1

14. There are no problems for a 2 year old child, already walking, to eat only lunch and dinner.

A child must have breakfast, lunch and dinner

1

15. It is enough for a child to eat fat rich foods 3 times a week.

5

Because those foods have fat which give strength to the child's body

16. A two year old child does not need the mother's supervision while eating.

1

The child needs supervision while eating

17. Sharing the plate or eating in his/her own plate is the same. In the end the child eats the same quantity of food.

1

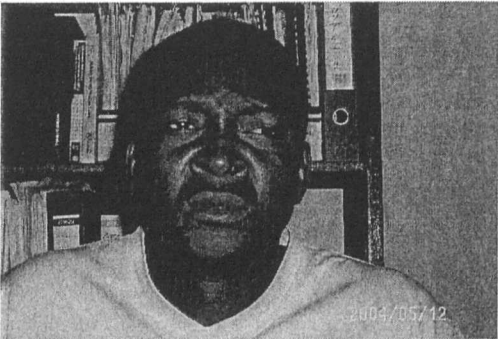
Because a child must eat in her own plate

18. The mother managed to get a bit of pig liver, but it is enough for everyone in the family, so the father will have it.

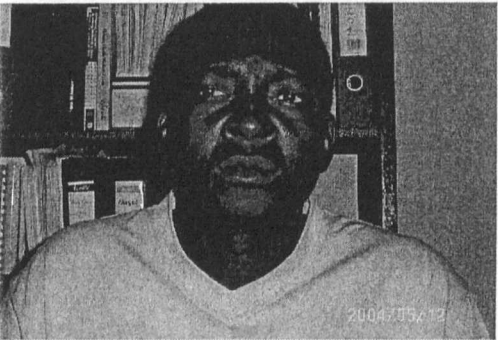
1

Liver is for the children

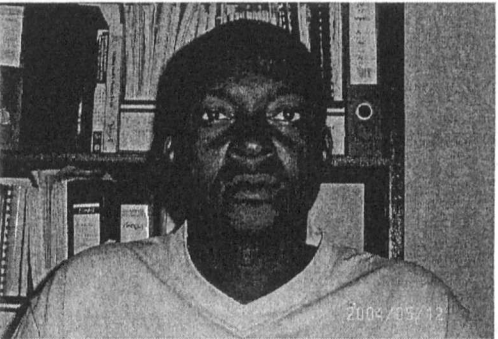
Appendix 4.5 – Faces used as the Likert scale



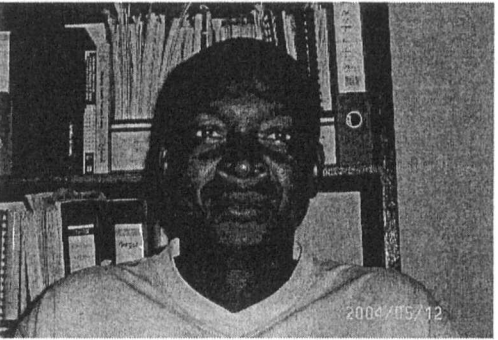
Strongly disagree



Disagree



Not sure



Agree



Strongly agree

Appendix 4.6 – Example of vignettes and corresponding story used in the attitude tool

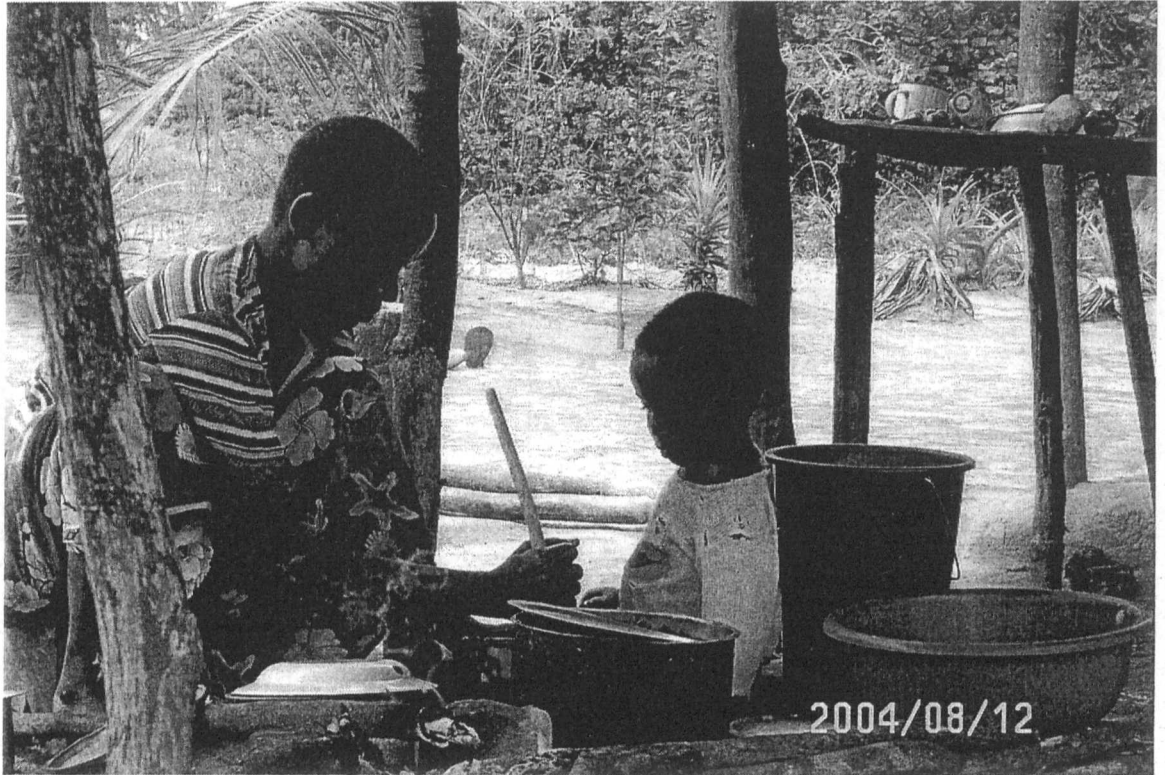
Item 7



Isabel and Teresa are sisters, and this is Isabel's baby which is 3 months old. Isabel is laid down because she has a strong malaria. She wanted to stop breastfeeding her baby but her sister Teresa convinced her not to do so. And that is why Isabel is still breastfeeding her child even though she has malaria.

Do you agree, disagree or are not sure with Isabel's decision to continue breastfeeding?

Item 11



Sarita has grown and she is now 3 years old. She is with her mother and both are preparing a porridge for Sarita's little sister that is only 6 months old. Sarita's mother has egg but she will not add it to the porridge because she thinks that a 6 months' baby cannot eat egg.

Do you agree, disagree or are not sure with Sarita's mother decision of not giving the egg to the baby?

Appendix 4.7 – Example of answers given when attitudes to complementary feeding items were negative

Groups	Remove the first milk and dispose of it	Give herbal tea to the new born baby	Decrease breast feeding and it will decrease diarrhoea	Stop breastfeeding at 6 months to be able to go to the field	Continue breastfeeding even if the mother is a month pregnant	Stop breastfeeding if the mother is 4 months pregnant	Stop breastfeeding a 9 month baby if the mother is ill with malaria	Stop breastfeeding if mother has not breastfed for 3 days
Intervention I	<p>"the first milk is bad for the infant and causes diseases"</p> <p>"need to remove it so the good milk can come"</p>	<p>"the tea is very good for the infant", "gives him/her health", "strength" and "cures diseases in the infant"</p>	<p>"the milk is bad for the infant", "it increases diarrhoea", "the milk could be spoiled", "if she reduces breastfeeding then diarrhoea will pass quicker"</p>	<p>"the child is old enough to stop breastfeeding", "for the mother to work well"</p>	<p>"a pregnant mother cannot breastfeed her child", "the milk will make the child ill"</p>	<p>"the belly is already too big", "can cause diseases to the child"</p>	<p>"the baby is big enough to stop breastfeeding" and "is already eating"</p>	<p>"she did not breastfeed for 3 days and the milk is already spoiled", "watery" and "can make the child ill" "with diarrhoea"</p>
Intervention II	<p>"the first milk is bad for the infant and causes diseases", it is contaminated", "need to remove it so the good milk can come"</p>	<p>"when a child is born, he/she should be given this tea", "the tea is very good for the infant", "gives him/her health", "strength" and "cures diseases in the infant and in the breast of the mother"</p>	<p>"the milk is bad for the infant", "it increases diarrhoea", "the milk could be spoiled", "if she reduces breastfeeding then diarrhoea will pass quicker"</p>	<p>"the child is old enough to stop breastfeeding", "for the mother to work well"</p>	<p>"a pregnant mother cannot breastfeed her child", "the milk will make the child ill"</p>	<p>"a pregnant women cannot breastfeed", the belly is already too big", "can cause diseases to the child", "the child will take long to grow"</p>	<p>"A sick mother cannot breastfeed her child otherwise the child will also get sick", "the child is big enough to stop breastfeeding"</p>	<p>"she did not breastfeed for 3 days and the milk is already spoiled", "contaminated", "watery" and "can make the child ill" "with diarrhoea"</p>

	<p>"the first milk is bad for the infant and causes diseases", "it gives diarrhoea to the baby".</p> <p>"Need to remove it so the child can be healthy".</p> <p>"Because that is what happens with the people here"</p>	<p>"The child has to take that tea in order to remove the diseases in his/her body and to avoid getting diarrhoea when starting to breastfeed".</p> <p>"Cures diseases in the infant and in the breast of the mother".</p> <p>"The tea is very good for the infant", "gives him/her health", and "strength".</p> <p>"because it is something that happens here"</p>	<p>"the milk is bad for the infant", "it increases diarrhoea", "the milk could be spoiled", "if she reduces breastfeeding then diarrhoea will pass quicker".</p> <p>"A sick child has difficulties in breastfeeding".</p> <p>"Because it happens here"</p>	<p>"the child is old enough to stop breastfeeding",</p>	<p>"a pregnant mother cannot breastfeed her child",</p> <p>"when a pregnant mother breastfeeds her child, transmits diseases to him/her",</p> <p>"the child can get contaminate d", "and can get stomach aches"</p>	<p>"a pregnant women cannot breastfeed", the belly is already too big", "can cause diseases to the child", "the mother is avoiding transmitting any diseases to her child"</p>	<p>"A sick mother cannot breastfeed her child otherwise the child will also get sick", "the child is big enough to stop breastfeeding"</p>	<p>"she did not breastfeed for 3 days and the milk is already spoiled",</p> <p>"contaminated", and "watery",</p> <p>"the child will get stomach aches"</p>
--	---	---	--	---	---	--	--	--